

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 15:16:12 : Search time 2912 Seconds

(without alignments)
3666.694 Million cell updates/sec

Title: US-09-990-415a-2

Perfect score: 1345
Sequence: 1 MESKMGELPLDINIQEPRMD.....PAMAIPLIMDTLEKKDFLK 261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=framer_p2n.model -DEV=rlp
-Q=/cgn2_1/USPTO.spool_P/US09990415/runat_16102003_073806_9438/app.query.fasta_1.455
-DB=GenEmbl -GFM=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Blasum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09990415 -CGN_1_1_3745 -runat_16102003_073806_9438 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEIOBURY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1345	100.0	1826	6 AX061199	AX061199 Sequence
2	1345	100.0	2702	6 BD127742	BD127742 Primer fo
3	1345	100.0	2702	9 AK074707	AK074707 Homo sapi
4	1345	100.0	2735	9 BC000124	BC000124 Homo sapi
5	1345	100.0	3047	6 AX188304	AX188304 Sequence
6	1318	98.0	984	6 AX364964	AX364964 Sequence
7	1286	95.6	2699	10 AF276997	AF276997 Rattus no
8	1283	95.4	2710	10 AF325262	AF325262 Mus muscu
9	1283	95.4	2870	10 BC012208	BC012208 Mus muscu
10	1089	81.0	1134	6 AX511171	AX511171 Sequence
11	1089	81.0	1818	9 AF327346	AF327346 Homo sapi
12	1089	81.0	2367	9 AK056915	AK056915 Homo sapi
13	1089	81.0	2236	5 BC044027	BC044027 Xenopus l
14	1076	80.0	2236	5 BC005743	BC005743 Mus muscu
15	1064	79.1	2986	10 S70011	S70011 Rattus sp.
16	1059	78.7	1204	10 BC005743	BC005743 Mus muscu
17	1059	78.3	2803	10 AF325260	AF325260 Mus muscu
18	1053	78.3	752	6 BD124792	BD124792 Primer fo
19	1053	78.3	752	6 BD126228	BD126228 Primer fo
20	1026	76.3	2766	6 AK022827	AK022827 Homo sapi
21	1026	76.3	2766	6 AK022827	AK022827 Homo sapi
22	1015	75.5	2894	9 BC020517	BC020517 Homo sapi
23	961.5	71.5	1084	6 AX457932	AX457932 Sequence
24	897.5	66.7	984	6 AX457933	AX457933 Sequence
25	897.5	66.7	984	6 AX511173	AX511173 Sequence
26	852.5	63.4	181086	9 AL133215	AL133215 Human DNA
27	798	59.3	2642	6 BD159719	BD159719 Primer fo
28	798	59.3	2642	9 AK022838	AK022838 Homo sapi
29	786.5	58.5	2062	5 BC048044	BC048044 Danio rer
30	765	56.9	741	6 BD148081	BD148081 Primer fo
31	756	56.2	1394	3 BT001832	BT001832 Drosophila
32	754	56.1	1316	3 AX089419	AX089419 Drosophila
33	753	56.0	1537	10 AF325261	AF325261 Mus muscu
34	753	56.0	1617	10 BC019808	BC019808 Mus muscu
35	750	55.8	1772	6 AX392949	AX392949 Sequence
36	750	55.8	2499	9 AF462052	AF462052 Homo sapi
37	750	55.8	2608	9 BC022091	BC022091 Homo sapi
38	712.5	53.0	1575	6 AX740470	AX740470 Sequence
39	711.5	52.9	255686	2 AC132957	AC132957 Mus muscu
40	705	52.4	252403	2 AC137016	AC137016 Rattus no
41	705	52.4	294396	2 AC105485	AC105485 Rattus no
42	705	52.4	347368	2 AC096620	AC096620 Rattus no
43	693	51.5	806	6 BD146776	BD146776 Primer fo
44	656	48.8	262886	6 AC121209	AC121209 Rattus no
45	650.5	48.4	229337	2 AC137328	AC137328 Rattus no

RESULT 1

ALIGNMENTS

AX061199
LOCUS AX061199 1826 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 46 from Patent WO0078953.
ACCESSION AX061199
VERSION AX061199.1 GI:12406335
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
1 Lal, P., Yang, J., Yue, H., Hillman, J.L., Tang, Y.T., Bandman, O.,
Burford, N., Baughn, M.R., Azimzai, Y., Lu, D.A., Au-Yang, J. and
Paterson, C.
Human transport proteins
Patent: WO 0078953-A 46 28-DEC-2000;
Incyte Genomics, Inc. (US)

TITLE
JOURNAL Incyte Genomics, Inc. (US)

FEATURES
SOURCE 1. 1826
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 415 a 519 c 512 g 380 t

ORIGIN

Alignment Scores:
Pred. No.: 2,33e-115 Length: 1826
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-990-415a-2 (1-261) x AX061199 (1-1826)

QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
DB 449 ATGGAAGCAAAATGGGTGATTCCTTTAGACATCAATCCAGAACTCCCTCGGGAC 508
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
DB 509 CAAAGTACTTCCCTGGGAGAGCCCGGCACCTTTTCACTGTACGATCTCGAAATCTG 568
QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAspTyrArgAlaGly 60
DB 569 CTGCTGTCGGGGGACAGCTGGAAGCTTCTGGAACATCTGCAGAACTACAGGGCCGCG 628
QY 61 ValValThrProGlyIleThrGluAspIleLeuTyrPargAlaLysTyrValTyrAspSer 80
DB 629 GTGGTGAACCCAGGAGATCAGCGAGGACAGCTGTGGAGGGCCAGTATGTATGACTCC 688
QY 81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
DB 689 GCCCTTCATCCGGACACAGGGGAGAGGTGCTCTGATGGCCCATGTCAGCCAGGTG 748
QY 101 ProMetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
DB 749 CCCATGAACATGACATCCTGCTGCTGCTCAATTTCAAGAGAGAACCCCAACCTGG 808
QY 121 ValPheTyrPglIntRPValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
DB 809 GTGTTCTGGCAGTGGGTGAATCATCTCTCAATGCCATGTGTAATCTCCACCGCGCT 868
QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
DB 869 GGTGACACTCCCATCTGCTGAGGACCTGGGACAGCTATGTGAGGCCACCACTGGA 928
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysIleLeuProPheValGly 180
DB 929 GCTGTGGCAGAGGGCCCTGGGACTCAANCCCTCAACAGACACCTGCCCTTGTGTGGC 988
QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200

DB 989 AGATTGTGCCCTTTGGACAGAGTGGACCTGCCCACTGCATCAACATCCCTGATGAGG 1048
QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyArgLeuGlyTyr 220
DB 1049 CAGAGAGAGCTGCAGAGTGGCATCCCGGTGCTGATGAGGACAGCTGCAGAGCTTGGCTAC 1108
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
DB 1109 TCGGTGACTGCAGCCACAGAGGAAATCTTCCACAGTGGTGAATTCAGAAATCTGCATGGCG 1168
QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysAspPheLeu 260
DB 1169 ATTCCTGCATGGCCATCCACCACTGATCATGACACTCTGAGAGAGAAAGACTTCTG 1228
QY 261 Lys 261
DB 1229 AAG 1231

RESULT 2
BD127742 2702 bp DNA linear PAT 18-SEP-2002
LOCUS BD127742
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127742
VERSION BD127742.1 GI:23222687
KEYWORDS JP 2002017375-A/3173.
SOURCE JP 2002017375-A/3173.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 2702)
Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, T.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002017375-A 3173 22-JAN-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/3173
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key

FEATURES
SOURCE FT CDS Location/Qualifiers
1. 2702
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 608 a 757 c 699 g 638 t

ORIGIN

Alignment Scores:
Pred. No.: 3-72e-115 Length: 2702
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-990-415a-2 (1-261) x BD127742 (1-2702)

QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
DB 70 ATGGAAGCAAAATGGGTGATTCCTTTAGACATCAATCCAGAACTCCCTCGGGAC 129

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 15:16:12 ; Search time 2912 Seconds
(without alignments)
3666.694 Million cell updates/sec

Title: US-09-990-415A-2
Perfect score: 1345
Sequence: 1 MESKMGELPDLINIQEPRWD.....PAMAIPLIMLEKKDFLK 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2688711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xl
-MODL=frame+2n.model -DEV=xl
-O/cgn2.1/USPTO_spool_p/US09990415/runat_16102003_073806_9438/app_query.fasta.1.455
-DB=genemb1 -QEXT=fastap -SUFFIX=rge -MINMATCH=0.1 -IOPTCL=0 -LOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human4.0.cdi -List=45
-DOALIGN=200 -THR.SCORE=DCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=fto -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09990415.ecgn.1.1.37.5 -runat.16102003.073806.9438 -NCPU=6 -ICPU=3
-NO.WMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPLOCK=100 -LONGIOS
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
GenEmb1:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.ov:*
22: em.pat:*
23: em.ph:*
24: em.pl:*
25: em.ro:*
26: em.sts:*
27: em.un:*
28: em.un:*

29: em.vl:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pin:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1345	100.0	1826	6 AX061199	AX061199 Sequence
2	1345	100.0	2702	6 BD127742	BD127742 Primer fo
3	1345	100.0	2702	9 AK074707	AK074707 Homo sapl
4	1345	100.0	2735	9 BC000124	BC000124 Homo sapl
5	1345	100.0	3047	6 AX188304	AX188304 Sequence
6	1318	98.0	984	6 AX364964	AX364964 Sequence
7	1286	95.6	2699	10 AF276997	AF276997 Rattus no
8	1283	95.4	2710	10 AF325262	AF325262 Mus muscu
9	1283	95.4	2870	10 BC012208	BC012208 Mus muscu
10	1089	81.0	1134	6 AX511171	AX511171 Sequence
11	1089	81.0	1818	9 AF327346	AF327346 Homo sapl
12	1089	81.0	1839	9 AK056915	AK056915 Homo sapl
13	1089	81.0	2267	6 AX302019	AX302019 Sequence
14	1076	80.0	2236	5 BC044027	BC044027 Xenopus l
15	1064	79.1	2986	10 S70011	S70011 Rattus sp.
16	1059	78.7	1204	10 BC005743	BC005743 Mus muscu
17	1059	78.7	2803	10 AF325260	AF325260 Mus muscu
18	1053	78.3	752	6 BD124792	BD124792 Primer fo
19	1053	78.3	752	6 BD126228	BD126228 Primer fo
20	1026	76.3	2766	6 BD157602	BD157602 Primer fo
21	1015	75.5	2894	9 AK022287	AK022287 Homo sapl
22	961.5	71.5	1084	6 AX457932	AX457932 Sequence
23	961.5	66.7	984	6 AX457933	AX457933 Sequence
24	897.5	66.7	984	6 AX511173	AX511173 Sequence
25	897.5	63.4	181086	9 AL133215	AL133215 Human DNA
26	852.5	59.3	2642	6 BD159719	BD159719 Primer fo
27	798	59.3	2642	6 AK022938	AK022938 Homo sapl
28	786.5	58.5	2062	5 BC048044	BC048044 Danto rer
29	765	56.9	741	6 BD148081	BD148081 Primer fo
30	756	56.2	1394	3 BT001832	BT001832 Drosophil
31	756	56.1	1316	3 AY089419	AY089419 Drosophil
32	753	56.0	1537	10 AF325261	AF325261 Mus muscu
33	753	56.0	1617	10 BC019608	BC019608 Mus muscu
34	750	55.8	1772	6 AX392949	AX392949 Sequence
35	750	55.8	2499	9 AF462052	AF462052 Homo sapl
36	750	55.8	2608	9 BC022091	BC022091 Homo sapl
37	750	55.8	2608	6 AX740470	AX740470 Sequence
38	712.5	53.0	1575	6 AC132957	AC132957 Mus muscu
39	711.5	52.9	255686	2 AC137016	AC137016 Rattus no
40	705	52.4	252403	2 AC137016	AC137016 Rattus no
41	705	52.4	294396	2 AC105485	AC105485 Rattus no
42	705	52.4	347368	2 AC096620	AC096620 Rattus no
43	693	51.5	806	6 BD146776	BD146776 Primer fo
44	656	48.8	262886	6 AC121209	AC121209 Rattus no
45	650.5	48.4	229937	2 AC137328	AC137328 Rattus no

RESULT 1

ALIGNMENTS

AX061199 1826 bp DNA linear PAT 22-JAN-2001
 LOCUS AX061199 Sequence 46 from Patent WO0078953.
 DEFINITION AX061199
 ACCESSION AX061199
 VERSION AX061199.1 GI:12406335
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1
 AUTHORS Lal, P., Yang, J., Yue, H., Hillman, J.L., Tang, Y.T., Bandman, O.,
 Butford, N., Baughn, M.R., Azimzal, Y., Lu, D.A., Au-Young, J. and
 Patterson, C.
 TITLE Human transport proteins
 JOURNAL Patent: WO 0078953-A 46-28-DEC-2000;
 Incyte Genomics, Inc. (US)
 FEATURES
 SOURCE 1..1826
 Location/Qualifiers
 BASE COUNT 415 a 519 c 512 g 380 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,336-115 Length: 1826
 Score: 1345.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 US-09-990-415a-2 (1-261) x AX061199 (1-1826)
 QY 1 MetGUsErLySMeTgLyGluLeuProLeuAspIleAsnIleGlnGluProArGTrpAsp 20
 DB 449 ATGGAACCAAAATGGTGGAATGGCTTTAGACATCAACATCCAGAACCTCGCGGAC 508
 QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgSnuLeu 40
 DB 509 CAAAGTACTTCTCGGGAGAGCCCGGACCTTTTCTACGTGACTGATCTCGAATCTG 568
 QY 41 LeuSerSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 DB 569 CTGCTGTCGGGGGACACAGCTGGAAGCTTCTCGAATCTGAGAACACTACAGCGCGC 628
 QY 61 ValValThrProGlyIleThrGluAspIleLeuThrParGalaLysTyrValTyrAspSer 80
 DB 629 GTGTGACCCAGGAGATCACCGAGACCGAGCTGTGGAGCGCCAGATATGTATGATCC 688
 QY 81 AlaPheHisProAspThrGlyGluLysValIleValIleGlyArgMetSerAlaGlnVal 100
 DB 689 GCCCTTCATCCGGACACAGGGAGAGAGGTGGCTCGATGGCCGATGTCAGCCAGGTG 748
 QY 101 ProMetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 DB 749 CCCATGAACATGACATGATGCTGCTGATGCTCATCTCTACAGGAACACCCCAACCTG 808
 QY 121 ValPheTyrGlnTyrPValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 DB 809 GTGTCTGCGAGTGGGTGAATCAGTCTTCAATGCAATGCTTACTACTCCACCGCAGT 868
 QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValIleSerAlaThrThrGly 160
 DB 869 GGTGACACTCCCATCAGTGTAGGAGAGCTGGGACAGCTTATGTAGTGCACCACTGGA 928
 QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysLysLeuProProLeuValGly 180
 DB 929 GCTGTGGCCAGCGCCCTGGAGACTCAAAATCCCTCAACAGACACTGCCCCCTGGTGGC 988
 QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnGlyIleAsnIleProLeuMetArg 200

DB 989 AGATTGTGCCCTTTCAGACAGTGGCAGCTGCCCACTGCATCAACATCCCTCATGAGG 1048
 QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyr 220
 DB 1049 CAGAGAGAGCTGCGAGGTGGCATCCCGGTGGTGTAGAGGAGCTGAGGCTTGCTTAC 1108
 QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAla 240
 DB 1109 TCGGTGACTGCGCCAGCAGGAGGAAATCTTCCAGGTGGTGTATTCANAGAAATGCTGGC 1168
 QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeu 260
 DB 1169 ATCTCTGCATGGCCATCCACACTGATCATGACATCTGTGAGAGAAAGACTTCTCG 1228
 QY 261 Lys 261
 DB 1229 AAG 1231
 RESULT 2
 BD127742 2702 bp DNA linear PAT 18-SEP-2002
 LOCUS BD127742
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD127742
 VERSION BD127742.1 GI:23222687
 KEYWORDS JP 2002017375-A/3173.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 2702)
 AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
 Makamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
 Koga, H.
 TITLE Primer for synthesizing full-length cDNA and use thereof
 JOURNAL Patent: JP 2002017375-A 3173 22-JAN-2002;
 COMMENT HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002017375-A/3173
 PD 22-JAN-2002
 PF 07-JUL-2000 JP 2000253172
 PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
 PI ISHII,
 PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
 SHINICHI KOJIMA,
 PI TETSUJI OTSUKI, HISASHI KOGA
 PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
 10,
 PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH key
 FEATURES
 source FT CDS Location/Qualifiers
 1..2702
 Location/Qualifiers
 BASE COUNT 608 a 757 c 699 g 638 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3-726-115 Length: 2702
 Score: 1345.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 US-09-990-415a-2 (1-261) x BD127742 (1-2702)
 QY 1 MetGUsErLySMeTgLyGluLeuProLeuAspIleAsnIleGlnGluProArGTrpAsp 20
 DB 70 ATGGAACCAAAATGGTGGAATGGCTTTAGACATCAACATCCAGAACCTCGCGGAC 129

QY 21 GlnSerThrPheLeuGlyAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
 DB 130 CAAAGTACTTCTTGGGAGAGCCGGCAGCTTTTACGTACTGATCTCGAATCTG 189
 QY 41 LeuLeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 DB 190 CTGCTGTCGGGGGACAGCTGAGACTCTTCGGAACATCGTCAGAACTACAGGGCGGCG 249
 QY 61 ValValThrProGlyIleThrGlnAspGlnLeuThrPheArgAlaGlyTyrValTyrAspSer 80
 DB 250 GTGGTGACCCCGGAGATCACCGAGACCGCTGTGAGGGCCCAAGTATGTGATGCTCC 309
 QY 81 AlaPheHisProAspThrGlyGlnValValIleLeuIleGlyArgMetSerAlaGlnVal 100
 DB 310 GCCTTCATCCGACACAGAGGAGAAAGGTGCTCGATGGCCGATGCGACCCAGG 369
 QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgIleThrProThrVal 120
 DB 370 CCCATGAACATGACCATCATCTGCTGCATGCTACATCTCTACAGAGAACCCCAACCGTG 429
 QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 DB 430 GTGTTCTGGCAGTGGGTGATCATGCTTCAATGCCATGTGTAACACTCCAAACGCACT 489
 QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
 DB 490 GGTGACATCCCATCATCTGTGAGGAGCGGTGGGAGACGCTATGTAGTGCACACCATGGA 549
 QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGly 180
 DB 550 GCTGTGGCCACGCGCCCTGGGACTCAATCCCTCACACMACACTGTGCCCTTGGTGCAG 609
 QY 181 AcgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
 DB 610 AATATTGTGCTCTTTCAGACAGTGGCAGCTGCCAAGTGCATCACTCCCTGATGAGG 669
 QY 201 GlnArgGlnLeuGlnValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyr 220
 DB 670 CAGAGAGAGCTGACAGTGGGATCCAGTGGCTGATGAGAGAGGTGACAGGCTTGGCTAC 729
 QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAla 240
 DB 730 TCGGTGACTGACAGCAGAGGAAATCTTCAGGTGTGATTTCAAGAACTGCAAGGGG 789
 QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeu 260
 DB 790 ATTCTGCGCATGGCCATCCACACATGATCATGAGACACTGTGAGAAAGAAAGCTTCTCG 849
 QY 261 Lys 261
 DB 850 AAG 852
 RESULT 3
 AK074707 2702 bp mRNA linear PRI 03-SEP-2002
 LOCUS Homo sapiens cDNA FLJ90226 f1s, clone NT2RM100858, weakly similar
 DEFINITION to tricarboxylate carrier.
 ACCESSION AK074707
 VERSION AK074707.1 GI:22760323
 KEYWORDS oligo capping; f1s (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS 1 Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,
 Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
 Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
 Kojima, S., Nagahara, K., Masuno, Y., Ono, T., Okano, K., Yoshikawa, Y.,
 Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
 Nishimura, K.
 TITLE NED0 human cDNA sequencing project

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2702)
 AUTHORS Isogai, T. and Otsuki, T.
 TITLE Direct Submission
 JOURNAL Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 COMMENT NED0 human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan: cDNA full insert sequencing:
 Research Association for Biotechnology; cDNA library construction:
 Institute of Medical Science, University of Tokyo, Laboratory of
 Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
 sequencing and clone selection: Helix Research Institute (supported
 by Japan Key Technology Center etc.).
 FEATURES
 source 1. 2702
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NT2RM100858"
 /cell_line="NT2"
 /cell_type="teratocarcinoma"
 /note="cloning vector: pUC19FL3-mRNA from uninduced NT2
 neuronal precursor cells"
 /note="unnamed protein product"
 /protein_id="BAC1151.1"
 /db_xref="GI:22760324"
 /translation="MESHKGEPLDINTEPRKDDSTFLGRARHFTYVDPRLILSG
 AOLEASRNIVQNRAGVVPFGTEDELMAKAYDSAFHPDGEKVVLLGRSAQVPM
 NMATIGCMLEFYRKPTPTVFMQVNSFNAYVSNRSDDTPITVQLGTAVASATG
 AVATLGLSLTKRLPLVGRVPEVFAVAANCINPLMKORELQVIGVADAGCRL
 GYSVTAAGKQIFQVYSRICMAIPMAVILMDIETKREKRLFRWLP84QVGLVG
 RCLVPTPLCCALFPQKSSIHISNLEPRLAQIHEDNPEVEVYVYKGL"
 BASE COUNT 608 a 757 c 699 g 638 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,72e-115 Length: 2702
 Score: 1345.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-990-415a-2 (1-261) x AK074707 (1-2702)
 QY 1 MetGlnSerLysMetGlnGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
 DB 70 ATGGAAGCAAAATGGGTGAATGCTTTAGACATCAACATCCAGGAACCTCGTGGAGC 129
 QY 21 GlnSerThrPheLeuGlyAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
 DB 130 CAAAGTACTTCTTGGGAGAGCCGGCAGCTTTTACGTACTGATCTCGAATCTG 189
 QY 41 LeuLeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 DB 190 CTGCTGTCGGGGGACAGCTGAGACTCTTCGGAACATCGTCAGAACTACAGGGCGGCG 249
 QY 61 ValValThrProGlyIleThrGlnAspGlnLeuThrPheArgAlaGlyTyrValTyrAspSer 80
 DB 250 GTGGTGACCCCGGAGATCACCGAGACCGCTGTGAGGGCCCAAGTATGTGATGCTCC 309
 QY 81 AlaPheHisProAspThrGlyGlnValValIleLeuIleGlyArgMetSerAlaGlnVal 100
 DB 310 GCCTTCATCCGACACAGAGGAGAAAGGTGCTCGATGGCCGATGCGACCCAGG 369
 QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgIleThrProThrVal 120
 DB 370 CCCATGAACATGACCATCATCTGCTGCATGCTACATCTCTACAGAGAACCCCAACCGTG 429

QY 121 ValPheTrpGlnTrpValAlaEngInSerPheAsnAlaIleValAsnTySerAsnArgSer 140
 DB 430 GTGTTGGCAGGGGGATGATCATCTCTTCAATGCCATTCTTAACACTCAACCCAGT 489
 QY 141 GLYAspThrProIleThrValArgGlnLeuGlyThrAlaTyValSerAlaThrThrGly 160
 DB 490 GGTGACACTCCATCTGATGAGCAGCTGGGACAGCCATATGATGAGTGCACACACTGA 549
 QY 161 AlaValAlaThrAlaLeuGlyLeuTySerLeuThrIleHisLeuProProLeuValGly 180
 DB 550 GCTGTGGCCAGGCGCCCTGGGAGCTCAATCCCTCACCAAGACATTCGCCCTTGTCGGC 609
 QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
 DB 610 AGATTGTGCCCTTTCAGACAGTGGCAGCTGCCAATCATCCATCCCTGATGAGG 669
 QY 201 GlnArgGlnLeuGlnValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTy 220
 DB 670 CAGAGAGAGCTGCAGGTGGCATCCAGTGGTGTGATGAGCAGGTGAGAGCTTGCTTAC 729
 QY 221 SerValThrAlaAlaValGlyGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
 DB 730 TCGGTGACTGCACCAAGCAGGAGATCTTCAGGTGTGATTCAGAAATCTGCATGCGC 789
 QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnIleTyIleAspPheLeu 260
 DB 790 ATTCTGCGCATGGCCATCCACCATCATATGATGACACTGTGGAGAAAGAACTTCCTG 849
 QY 261 Lys 261
 DB 850 MAG 852

RESULT 4
 LOCUS BC000124 2735 bp mRNA linear PRI 12-JUL-2001
 DEFINITION Homo sapiens, similar to tricarboxylate carrier-like protein, clone
 ACCESSION MGC:3122 IMAGE:3352015, mRNA, complete cds.
 VERSION BC000124
 KEYWORDS BC000124.1 GI:12652746
 SOURCE MGC.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 2735)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (03-NOV-2000) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 CONTACT: MGC help desk
 EMAIL: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 Info@gscc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Steven
 Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 6 Row: K Column: 13
 This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF
 analysis, Genomescan gene prediction.
 Location/Qualifiers

FEATURES

source
 1. 2735
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:3122 IMAGE:3352015"
 /tissue_type="Eye, retinoblastoma"
 /clone_id="NIH_MGC_16"
 /lab_host="DH10B-R"
 /note="Vector: pOTB"
 75..1040
 /codon_start=1
 /product="Similar to tricarboxylate carrier-like protein"
 /protein_id="AAH00124.1"
 /db_xref="GI:12652747"
 /translation="MEGLPDINIOEPMDOSFELGRARHFTVTPRNLLSGAOE
 ASRNITQNRACGVYTTTEDQMRKYYTDSFHPDTEKVVLLGRMSAOVPMNTI
 TSCMLTFRTPTVFWQVNOSEFNAIVNSNDSGPTPIVROLGTAYSATGAVAT
 ALGLSLTRHLPLVGRFVFAVAANCLINPLMORLEOVGPVADDEGRLGYSV
 TAKOGIFQVVISRICMAIPAMAIPLIMDTLEKDFLKRPMGLGVLGVFCILY
 FATPLCCALFPQKSIHISNLEPELRAQIHQONPSVEVYVYNNKGL"

CDS

BASE COUNT 645 a 757 c 695 g 638 t
 ORIGIN

Alignment Scores:

Pred. No.: 3,78e-115 Length: 2735
 Score: 1345.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-990-415a-2 (1-261) x BC000124 (1-2735)

QY 1 MetGluSerIleuMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyPasp 20
 DB 63 ATGGAAAGCAAAATGGGTGAATTCCTTAGACATCAACATCCAGGAACCTCGCTGGAC 122
 QY 21 GlnSerThrPheLeuGlnArgAlaArgHisPheThrValThrAspProArgAsnLeu 40
 DB 123 CAAAGTACTTCTCTGGGAGAGCCGGCACTTTTACAGTCTTACTGATCTCGAAATCTG 182
 QY 41 LeuLeuSerGlyValAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyArgAlaGly 60
 DB 183 CTGCTGTCCGGGGCAGACGCTGGAAGCTTCTCGGAACATGTCGAGAACTACAGGCCGCG 242
 QY 61 ValValThrProGlyIleThrGlnAspGlnLeuThrArgAlaIleTyThrValTyAspSer 80
 DB 243 GTGGTACCCCAAGGATCACCGAGCAACCTGTGGAGGGCCAAAGTATGTATGACTCC 302
 QY 81 AlaPheHisProAspThrGlyGlyValValLeuIleGlyArgMetSerAlaGlnVal 100
 DB 303 GCCTTCATCCGGACACAGGGAGAGGTGCTCATTTGGCCGATGTCAGCCAGG 362
 QY 101 PrometAspMetThrIleThrGlyCysMetLeuThrPheTyArgIleThrProThrVal 120
 DB 363 CCCATGAATACATGACATCTGAGCTGACAGCTCACTTCTACGGAAGACCCCAACCGTG 422
 QY 121 ValPheTrpGlnTrpValAlaEngInSerPheAsnAlaIleValAsnTySerAsnArgSer 140
 DB 423 GTGTTTGGCAGGGGTGAATCATCTTCAATGCCATTGTTAACACTCAACCCAGT 482
 QY 141 GLYAspThrProIleThrValArgGlnLeuGlyThrAlaTyValSerAlaThrThrGly 160
 DB 483 GGTGACACTCCATCTGATGAGCAGCTGGGAGAGCCATATGATGAGTGCACCACTGA 542
 QY 161 AlaValAlaThrAlaLeuGlyLeuTySerLeuThrIleHisLeuProProLeuValGly 180
 DB 543 GCTGTGGCCAGGCGCCCTGGGAGTCAATTCCTCACCAAGACCTGCCCCCTTGTCGGC 602
 QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200

```

|||||
Db 603 AGATTGTGCGCTTGGACGACGAGTGGACAGTCCCAACTGCATCAACATCCCGCTGATGAGG 662
Qy 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluIleGlnArgLeuGlyTyr 220
Db 663 CAGAGAGAGCTGCGAGGTGGGATCCCGGTGCTGATGAGGATGAGGCTTGGCTTAC 722
Qy 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db 723 TCGGTACTGCGACGCAAGAGGAAATCTTCCAGGTGGTGATTTCAAGATCGATGCGG 782
Qy 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
Db 783 ATTCCTCCATGGCCATCCACACAGTATGACACTCTGAGAGAAAGAACTTCTCTG 842
Qy 261 Lys 261
Db 843 AAG 845

RESULT 5
AX188304/c 3047 bp DNA linear PAT 06-AUG-2001
LOCUS Sequence 3999 from Patent WO0142467.
DEFINITION AX188304
ACCESSION AX188304
VERSION AX188304.1 GI:15139777
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 Schlegel R., Deeds J., Berger A. and Zhao X.
Genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of cervical cancer
Patent: WO 0142467-A 3999 14-JUN-2001;
JOURNAL Millennium Predictive Medicine, Inc. (US)
FEATURES
Source Location/Qualifiers
1..3047
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 629 a 790 c 891 g 733 t 4 others
ORIGIN
Alignment Scores:
Pred. No.: 4.3e-115 Length: 3047
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-990-415a-2 (1-261) x AX188304 (1-3047)
Qy 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
Db 2605 ATGGAAGCAAAAGGATGATTCCTTACACATCAACATCCAGAACTCCGCTGGGAC 2546
Qy 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValIleTrpAspProArgAsnLeu 40
Db 2545 CAAGTACTTCTCGGGGACAGCCGCGACCTTTTCACTGTTACTGATCCCGAANAATCTG 2486
Qy 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
Db 2485 CTGCTGTCGGGGGACAGCTGGAAGCTTCTCGAACAATCGTGAAGAACTACAGGGCCGGC 2426
Qy 61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80
Db 2425 GTGGTGAACCCAGGAGATCCAGGAGACCACTGTGGAGGCCAAGTATGTATGACTCC 2366
Qy 81 AlaPheHisProAspThrGlyLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db 2365 GCCTTCATCCGGACAGAGGGGAGAGAGTGTCTGTGATGGCCGCAATCTCAGCCAGGTG 2306

```

```

Qy 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db 2305 CCCATGAACATGACATCTACTGCTGATGCTGATCTACATCTCTACAGGAACCCCAACCTG 2246
Qy 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db 2245 GTGTTCTGCGAGTGGGTGAATCAGTCCTTCAATGCTCATTTGTAATCTTCCAAACCGCAGT 2186
Qy 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrCly 160
Db 2185 GGTACACTCCCATCAGTGTAGGACGCTGGGAGAGCCTTATGTAGTGCACACACTGGA 2126
Qy 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysIleLeuProProLeuValGly 180
Db 2125 GCTGTGGCGACAGGCGCTGGAGACTCAAACTCCACCAAGACACTGCCCTTGTGCGC 2066
Qy 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
Db 2065 AGATTGTGCGCTTGGACAGTGGACACTGCCAACTGCATCAACATCCCTGATGAGG 2006
Qy 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluIleGlnArgLeuGlyTyr 220
Db 2005 CAGAGAGAGCTGCGAGGTGGCATCCCGGTGCTGATGAGGAGTCAAGAGCTTGGCTTAC 1946
Qy 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db 1945 TCGGTACTGCGACGCAAGAGGAAATCTTCCAGGTGGTGATTTCAAGATCTGCATGGCG 1886
Qy 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
Db 1885 ATTCCTCCATGGCCATCCACACAGTATGACACTCTGAGAGAAAGAACTTCTCTG 1826
Qy 261 Lys 261
Db 1825 AAG 1823

RESULT 6
AX364964 984 bp DNA linear PAT 15-FEB-2002
LOCUS Sequence 115 from Patent WO0206315.
DEFINITION AX364964
ACCESSION AX364964
VERSION AX364964.1 GI:18696854
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 Mintz L., Freilich S. and Bernstein J.
Novel nucleic acid and amino acid sequences
Patent: WO 0206315-A 115 24-JAN-2002;
JOURNAL Compugen Ltd. (IL)
FEATURES
Source Location/Qualifiers
1..984
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 224 a 278 c 285 g 196 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3.59e-113 Length: 984
Score: 1318.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 97.99% Indels: 0
Gaps: 0

US-09-990-415a-2 (1-261) x AX364964 (1-984)
Qy 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20

```

Db 49 ATGGAAGCAAAATGGTGAATTCCTTAGACATCAATCCAGGAACCTCGTGGAC 108
 QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThraSPProArgAsnLeu 40
 Db 109 CAATGCTACTTCTCTGGCAGAGC -CGGCACTTTTCTACTGTTACTGATCTCTGAAATCTG 167
 QY 41 LeuLeuSerGlyAlaGlnLeuGlnLalaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 Db 168 CTGCTGCTCCGGGGGACACAGCTGGAACCTTCGGAAACATCTGTCAGAACTACAGGGCCGGC 227
 QY 61 ValValThrProGlyIleThrGlnAspGlnLeuThrParGalaValTyrValTyrAspSer 80
 Db 228 GTGGTGACCCAGGATACCGAGAGCAGCTGTGGAGGGCCCAAGTATGTATACATCC 287
 QY 81 AlaPheHisProAspThrGlyGlnLysValValLeuIleGlyArgMetSerAlaGlnVal 100
 Db 288 GCTTCCATCCGGACACAGGGAGAGGTGGTCCGATGGATGGCCGATGACCCAGGCTG 347
 QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 Db 348 CCATGACATGACATCAGTCCGTCGATGCTCCTACATCTTACAGGAAGACCCCAACCGTG 407
 QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 Db 408 GTGTTCTGGCAGTGGGTGAATCAGTCTCAATGCCATTGTAATCTCCAAACCGCAGT 467
 QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
 Db 468 GGTGACACTCCCATCCTGCTGAGAGCAGCTGGGAGAGCCCTATGTAGTCCACCACTGGA 527
 QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
 Db 528 GCTGTGGCCAGCGCCCTGGAGCTCAATCCCTCAACCAAGCAACCTGCCCTTGGTCCGC 587
 QY 181 ArgPheValProPheAlaIleValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
 Db 588 AGATTGTGCTGCTTGGCAGCAGTGGCAGCTCCCAACTCACAATCCCTCCCTGTAGAG 647
 QY 201 GlnArgGlnLeuGlnValGlyIleProValAlaAspGlnValGlnArgLeuGlyTyr 220
 Db 648 CAGAGAGAGCTGAGGTGGCATCCGGTGGCTGATGAGCAGGTCAAGAGCTTGGCTTAC 707
 QY 221 SerValThrAlaAlaValGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
 Db 708 TCGGTGACTGCGACGACAGGAGATCTTCCAGGTGGATTCAGAAATCTGCATGGCG 767
 QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeu 260
 Db 768 ATTCTGCGATGGCATCCACCACTGATCATGACACTGTGAGAAAGAAAGACTTCTCTG 827
 QY 261 Lys 261
 Db 828 AAG 830
 RESULT 7
 AF276997 2699 bp mRNA linear ROD 22-JUN-2000
 LOCUS Rattus norvegicus tricarboxylate carrier-like protein mRNA,
 DEFINITION complete cds.
 ACCESSION AF276997
 VERSION AF276997.1 GI:8650525
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 2699)
 MASHIMA,H. and KOJIMA,I.
 Direct Submission
 Submitted (12-JUN-2000) Cell Biology, Institute for Molecular and
 Cellular Regulation, Gunma University, 3-39-15, Showa-machi,
 Maebashi, Gunma 371-8512, Japan

FEATURES
 source
 Location/Qualifiers
 1..2699
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague Dawley"
 /db_xref="taxon:10116"
 233..1198
 /note="CTP-like protein"
 /product="tricarboxylate carrier-like protein"
 /protein_id="AF78249.1"
 /db_xref="GI:8650526"
 /translation="MGDLPLNINIDPEKMDSTFGRARHFTVTDPNNLSIKQLE
 ASRNIVQNTKRVAGVTPFGLTEQIMRAKYVDSAPDPEKVVLLGRMSAOVPMNTI
 TGM/LTFKRTPTVFMQVNSFNAIVNSRSDAPITVOQLTAYASAVAT
 ALGLSLTKHLPLVGRFVPAVAANICINPLRORLELDVGIPTDEAGORLGSV
 TAAKGIFQVNVSRGMIIPAAIPIPVNMTLEKDFLKRPMGLAPQVGLVGRCLV
 FAPLPCALFPQRRSIHNTRELELRADIQAKPSIDVYVYKGL"
 2661..2666
 polyA_signal
 BASE COUNT 611 a 747 c 684 g 657 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1
 Score: 1286.00
 Percent Similarity: 98.46%
 Best Local Similarity: 95.38%
 Query Match: 95.61%
 DB: 10
 Gaps: 0
 US-09-990-415a-2 (1-261) x AF276997 (1-2699)
 QY 2 GlnSerLysMetGlyGlnLeuProLeuAspIleAsnIleGlnGlnProArgTrpAspGln 21
 Db 224 GAAGCAAGATGGGGGAGCTTCCCTTAATATCAACATCCAGGAACCTCGATGGGACCA 283
 QY 22 SerThrPheLeuGlyArgAlaArgHisPhePheThrValThraSPProArgAsnLeu 41
 Db 284 AGCATCTTCTTAGGACAGAGCCCGGACCTTTCACAGTTACTGATCCCGAAATCTCTG 343
 QY 42 LeuSerGlyAlaGlnLeuGlnLalaSerArgAsnIleValGlnAsnTyrArgAlaGly 61
 Db 344 CTGTCGGGGAACAGTTGGAAGCTTCTCGGAACATCGTGAAGATTCAGAGGCTGTGTG 403
 QY 62 ValThrProGlyIleThrGlnAspGlnLeuThrParGalaValTyrValTyrAspSerAla 81
 Db 404 GTAACCTCTGGGCTCAGTGAAGCAGCTATGAGAGCCAAATAGCTGTATGACTACACA 463
 QY 82 PheHisProAspThrGlyGlnLysValValLeuIleGlyArgMetSerAlaGlnValPro 101
 Db 464 TTCCACCCGGACACAGGGAGAGAGTGTCTTGATGTGTCGCATGTCCGCCAGGTGCC 523
 QY 102 MetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValVal 121
 Db 524 ATGAACATGACCATTAAGCTGCTGCATGCTCCTACAGAAAGACTCAACTGTGGTG 583
 QY 122 PheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerArgSerGly 141
 Db 584 TTGTGGCGAGTGGGAATCAGTCTCAATCTATGATGATTAATCGCAGTGA 643
 QY 142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAla 161
 Db 644 GAGCTCCTATCAGTGCAGACAGCTGGGAACGGCTATGAGAGCCACACAGTGGGCT 703
 QY 162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArg 181
 Db 704 GTGGCTACTGCTTGGGACTCAAGTCTTCCCAAGCAACCTACCCCTGTAGGTGCGA 763
 QY 182 PheValProPheAlaIleValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGln 201
 Db 764 TTCTACCTTTGACAGCTGTGGTGGCCCAACTGCATCAACATCCCTTATGAGGCGAG 823
 QY 202 ArgGlnLeuGlnValGlyIleProValAlaAspGlnValGlnArgLeuGlyTyrSer 221

Db	824	AGGAGAGCTGCACGGTGGCCATCCCATGTCACTACTGACAGCTGGTCAGACACTTGCCCATCG	883
Oy	222	ValThrAlaAlaIatrygInglyLilephedInvalVallieserrArgIllecysMetAlalle	241
Db	884	GTCAGCTGTCGCCAACAACAGGCATCTCCAGGTGGGGATGATCGAGAATCGCATGGCGCATC	943
Oy	242	ProAlaMetaIalleleProIoleuIleMetAsPthrIleugIlyslYlsAspPhelauDys	261
Db	944	CCCCGCATGCGCATTCGCCCGGTGATCATGAACAACACTCTGGAGAAAGAACTTCCTCTCAG	1003
RESULT 8			
AF325262		2710 bp	mRNA linear ROD 30-APR-2001
LOCUS	Mus musculus sleroflexin 3 (Sfxn3) mRNA, complete cds.		
DEFINITION	AF325262		
VERSION	AF325262.1 GI:13785615		
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	Fleming, M.D., Campagna, D.R., Haslett, J.N., Trenor, C.C. III and Andrews, N.C.		
JOURNAL	A mutation in a mitochondrial transmembrane protein is responsible		
MEDLINE	for the pleiotropic hematological and skeletal phenotype of		
PUBMED	flexed-tail (f/f) mice		
REFERENCE	Genes Dev. 15 (6), 652-657 (2001)		
AUTHORS	2 (bases 1 to 2710)		
TITLE	Fleming, M.D., Campagna, D.R. and Andrews, N.C.		
JOURNML	Submitted (01-DEC-2000) Pathology, Children's Hospital, 300		
FEATURES	Longwood location/Qualifiers		
SOURCE	1..2710		
gene	/organism="Mus musculus"		
CDS	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
	1..2710		
	/gene="Sfxn3"		
	83..1048		
	/gene="Sfxn3"		
	/note="SFXN3, similar to the flexed-tail (f) protein		
	sleroflexin 1 (SFXN1)".		
	/codon_start=1		
	/product="sleroflexin 3"		
	/protein_id="AAK39430.1"		
	/db_xref="GI:13785615"		
	/translation="MGDLPLINIQEPRMDSTFLGARHFEFTVDPRNLLSGEQIE		
	ASRNIVONYRAGVATPGITEDOLARAYVDASHPTGEAVYLIGMSAOVMNNITI		
	TSCMLTYERKTPYTFQWOWNSFNALIVNSNRSGDAPIVOQLGTAVSATGVANN		
	ALGKSLKLHLPLVLGRFVPFAVAALNCISIPLMRORELOYGIPVDEGORLGHSV		
	TAAOGSEIOVISRIEMAIIPMAIPVIMNTLEKKDKLPRTGAPLVQGVNGFCFLV		
	PATPLCCALPFQRSIHVTRLEPELRLOIAQNPSIDVVYINKGL"		
BASE COUNT	591 a 755 c 647 g 711 t	6 others	
ORIGIN			
Alignment Scores:			
Pred. No.:	2,14e-109	Length:	2710
Score:	1283.00	Matches:	248
Percent Similarity:	98.08%	Conservative:	7
Best Local Similarity:	95.38%	Mismatches:	5
Query Match:	95.39%	Indels:	0
DB:	10	Gaps:	0
US-09-990-415A-2 (1-261) x AF325262 (1-2710)			
Oy	2	GlusterlysmetGlygluleuProleuaspIleasnillegingluiproArtyTPPaPoln	21
Db	74	GAAGGCAAGATGGGTGACCTCCCTTAATATCAACATCCAGGAACCTCGGTGGACCNA	133

OY		22	SerThrPheLeuGlyArgAlaArGIshPheSerThrValTTrAspProArGIshLeuLeu	41
Dd		134	AGCACATTCTTAGGCACAGACCAGCCGGCATTTCTTCACAGCTACGTACTCGAATCTGGCTG	193
OY		42	LeuSerGIyAlaGlnLeuGlnLysAlaSerArgAsnIleValIGlnAsnTYrArgAlaGlyVal	61
Dd		194	CTGTCCGGGGGAACAGCAGTGAAAGCTTCCCGGAACATCGTGAAGTAATTAACAGGGCTGTGTG	253
OY		62	ValThrProGlyIleThrGlnAspGlnLeuThrPargAlaLysTyValTYrAspSerAla	81
Dd		254	GCAACCCCGGGGTCTTCACAGAGACACAGTATGAGCGACAATACTGTATGACTCACACA	313
OY		82	PheHisPProAspThrGlyILysValValLeuIleGlyArgMetSerAlaGlnValPro	101
Dd		314	TTCATCTCCGGACACAGGGGAGAAGGTGCTTCTTGATGTGCCCGTATGTACAGCCAGGTGCC	373
OY		102	MetAsnMetThrIleThrGlyCyseMetLeuThrPheTYrArgLYsThrProThrValVal	121
Dd		374	ATGAACATGACCATTAAGTGGCTGATGCTACCTTCACAGGAAGACATCCGACTGTGGTG	433
OY		122	PheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTYrSerAsnArgSercly	141
Dd		434	TTCGTGGCAGTGGGTCAATCAGTCCCTTCATCTATTGTGAATTTACTTCAATCGCAGCGCG	493
OY		142	AsphTrpProIleThrValArgGlnLeuGlnLythrAlaTYrValSerAlaIleThrGlyAla	161
Dd		494	GATGCTCCCATCAGCTGTGCAGAGTGGGGACAGCTATGTGAGTGGCCACCATGTGGGCT	553
OY		162	ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArg	181
Dd		554	GTGGCTACTGCTGTGGGACTCAAGTCTCTACCAAAGACATCGCCCGCTAGTGGTGA	613
OY		182	PheValProPheAlaIleValAlaAlaAsnCysIleAsnIleProLeuMetArgGln	201
Dd		614	TTCTGTGCCCCCTTGGCAGCTGTGGCGCTGCCAACGTCAATCAATCCCCTATATAGCGAG	673
OY		202	ArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyArgLeuGlyTYrSer	221
Dd		674	AGGAGACTCGCAGGTGGGCATCCCACTACTATGATAGGCTGTACAGAGCTTGGCCACTCG	733
OY		222	ValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgGlyIleCyseMetAlaIle	241
Dd		734	GTGACTGCTGCCAACAAGAGAAATCTTCCAGGTGGGATATCAAGAAATCGAATGGCGATT	793
OY		242	ProAlaMetAlaIleProProLeuIleMetAspHrLeuGlnLysLysAspPheLeuLys	261
Dd		794	CCCCGCCAAGCCATTCCCCCGGTGATCATGAACACTGTGGAGAAAGAAAGACTTCTCTGAG	853
LOCUS				
DEFINITION				
ACCESSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				

TITLE	Proteins
JOURNAL	Patent: WO 0246221-A 1 13-JUN-2002; Oxford GlycoSciences (UK) Limited (GB)
FEATURES	location/Qualifiers
source	1..1134

BASE COUNT	303 a	268 c	273 g	287 t	3 others
ORIGIN					

Alignment Scores:

Pred. No.:	7,85e-92	Length:	1134
Score:	1089.00	Matches:	201
Percent Similarity:	89.45%	Conservative:	28
Best Local Similarity:	78.53%	Mismatches:	27
Query Match:	80.97%	Indels:	0
DB:	6	Gaps:	0

US-09-990-415A-2 (1-261) x AX511171 (1-1134)

QY	6	GlyGlnLeuProLeuAsnSprLLeAsnInLeGlnInPuProAqTTPAsGInSerThPhLeu	25
Db	17	GGAGAACTACACCAACATTAACATCAAGAAAGACTCCATGGATCAAGACATTTCAATT	76
QY	26	GlyAlaGAlaIaArgHisPhePheTherValThirAspProArgAsnLeuLeuSerGlyAla	45
Db	77	GGACGGAGCCAAATCATTTCTTCACTGTAAGTACGCCCAAGAAACATCTGTTAACCAAGAA	138
QY	46	GlnLeuGlnAlaSerArgAsnLLeValGlnAsnTyTrArgAlaGlyAlaValThProGly	65
Db	137	CAACTCGAAGAGCGCGAGAAAATTAATGATACAGTGTACAGGCAAGAAATGTTCCCTCGGT	196
QY	66	IleThrGluAspGlnLeuThrParArgAlaLysTyTrValTyTrAspSerAlaPheHisProAsp	85
Db	197	CTTACAGAAATAGAAATGTGTGGAGACCAAGTACATCTGTGATTCAGCTTTTCACTCGTAC	255
QY	86	ThrGlyGluLysValAlaLeuIleGlyArgMetSerAlaGlnValPrometAsnMetThr	105
Db	257	ACTGGTGAAGAAATGCTTTTGTATGAGGAAGATGTACGCCAGGTGCCATGACATGACC	318
QY	106	IleThrGlyCyMetLeuThrPhePheTyArgTyrThrProThrValAlaPheTrgInTTrp	122
Db	317	ATCACAGGTGTATGATGTACGTTTACAGGACATACGCGGCTGTGCTGTCTTGCCAACTGG	378
QY	126	ValAsnGlnSerPheAsnAlaIleValAsnTyTrSerAsnArgSerGlyAspThrProIle	145
Db	377	ATTAAACCAATCCTTCAATGCCCTGCTCAATTACACCAACAGATGGAGACCAACCCCTC	433
QY	146	ThrValArgGlnLeuGlyThrAlaTyTrValSerAlaThrThrGlyAlaValAlaThrAla	165
Db	437	ACTGTCAATAGATGGGAACACAGCTTACAGTTTCTGCAACAATGGTGGCGGTACCAAGCT	498
QY	166	LeuGlyLeuLysSerLeuThrLysHisLeuProPoleuValGlyArgPheValProPhe	185
Db	497	CTAGGACTCAATGCATGTACCAAGAGATGCTCACACAGATGAAGAGAGCTTTGTCCCTTT	556
QY	186	AlaAlaValAlaAlaAlaAsnCysLlleAsnIleProLeuMetArgGlnArgGlnLeuGln	205
Db	557	GCTGCCGTAGCTGCTGCTATATGGCAATTATATCTCATTAAGAGGCAAAAGGAACCTCTAAA	618
QY	206	ValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyTrSerValThrAlaAla	225
Db	617	GTTGGGATTCCTCGTACGGATGAGAATGGAAACCCCTTGGGGAGATCGCGCAACGCTGGC	678
QY	226	LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla	245
Db	677	AAACAAGCAATACAGCAAGATGCTGTCGTCACAGGATTCATCAGCAAGCCCTCGCATGGCC	738
QY	246	IleProPoleuIleMetAspThrLeuGlnLysLysAspPheLeuLys	261
Db	737	ATCCCTTCATTCATTATGAACACTTGTGGAAAAAGAAAGCCCTTTTGAAG	784

RESULT 11	
AF327346	
LOCUS	1818 bp mRNA linear
DEFINITION	Homo sapiens tricarboxylate carrier protein mRNA, complete cds.
ACCESSION	AF327346
VERSION	AF327346.1 GI:18028274

FEATURES
source

CDS

```

                                /organism="Homo sapiens"
                                /mol_type="mRNA"
                                /db_xref="taxon:9606"
                                62. .1030
                                /codon_start=1
                                /product="tricarboxylate carrier protein"
                                /protein_id="AA156007.1"
                                /db_xref="GI:18028275"
                                /translation="MSGELPENINIKERPMQSTFIGRANHFVTVDPRNILLTNEQI
                                ESARKIVDHYRQGIYPPGLTENELMRKYYIDSFAPHDTGKMLIGMSAQVPMNMT
                                ITGCMFTFRTTPAVLPMQMINOSFNVNTNSGAPLVNDELGRVYATGGAIV
                                TAGLNLTLKHSVPLIGRFVPEFAVAANCINTPLMRQELKVGIPDDEGNRLEE
                                ANAKOALTOYVYSRIIMADGALIPPIINMTLEKKAFLKRFPMMSAPIQVGLVGFCCI
                                VFATPICALFPQKSSMSVTSLEELAKIDSEAPELRRRYFNKGL"
BASE COUNT      551 a      375 c      403 g      469 t

```

BASE COUNT	551 a	375 c	403 g	489 t
ORIGIN				
Alignment Scores:				
Pred. No.:	1.38e-91			1818
Score:	1089.00			Matches: 201
Percent Similarity:	89.45%			Conservative: 28
Best Local Similarity:	78.57%			Mismatches: 27
Query Match:	80.97%			Indels: 0
DB:	9			Gaps: 0

US-09-990-415A-2 (1-261) X AF327346 (1-1818)

QY	6	GlyGluLeuProLeuAspIleAsnIleGlnGluProAlaGlyThrPheGlnSerThrPheLeu	25
	68	GGGAAGACTCCACCCAACTTAACATCAAGGAACCTCATGGATCAAAAGACCTTCAAT	127
Db	26	GlyArgAlaArgHisPhePheThrValThrAspProGlnAsnLeuLeuSerGlyAla	45
	128	GGACGAGCCAAATCATTTCTTCTACTGTAACTACCCCAAGAACATTTCTTTAACCAACGAA	187
QY	46	GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly	65
	188	CAACTCGAAGTGGCGAAGAAATAGTACATGATTACAGCAAGGAATGTCTCCCTGGT	247
Db	66	IleThrGluAspGlnLeuThrPargAlaLysTyrValTyrAspSerAlaPheHisProAsp	85
	248	CTTACCGAAAGAAATGTGGTGGAGACCAAGTACATCATGTGATTCAGCTTTTCAATCCGAC	307
QY	86	ThrGluGluIleValLeuIleGlyArgMetSerAlaGlnValProMetAsnMetThr	109
	308	ACTGGTGAAGAGTGTGTTTGTATGAGAAAGATGTAGCCCAAGTTCCTCATTAACATGACC	367
Db	106	IleThrGlyCysMetLeuThrPheTyrArgTyrThrProThrValValPheThrGlnTyr	125
	368	ATCACAGGTGTATGATGATGACGTTTAAACAGACTACGCCGGCGTGGCTTCTGTGGCAGTGG	427
QY	126	ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle	145
	428	ATTAAACAGTCCCTTCATCCCTCGCTCAATTTACACCAACAGAAAGGAGAGCACCCCTC	487
Db			

QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyValAlaValAlaThrAla 165
 Db 488 ACTGTCATGAGTGTGGAAACACCTTACGTTTCTGCAACACCTGGCCCTACCAACAGCT 547
 QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
 Db 548 CTAGAGCATCAATGACATGACCAAGCATGCTCCACACGATGATGAGAGCTTTGTCCCTT 607
 QY 186 AlaAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 205
 Db 608 GCTGCCGTAAGTCTGCTGCTAATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 667
 QY 206 ValGlyIleProValAlaAlaAspGluAlaGlyIleArgLeuGlyTyrSerValThrAlaAla 225
 Db 668 GTTGGCATTTCCCGTCACGATGAGATGGAACCGCTTGGGGAGTCCGCGCAACCTCGC 727
 QY 226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
 Db 728 AACACAGCCATCACCAAGTGTGCTGTCAGAGATTCATGCGCAGCCCTGGCATGGCC 787
 QY 246 IleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeuLys 261
 Db 788 ATCCCTCATTCATTATGACACTTTGGAAGAAAGAGCCCTTTTGAAG 835
 RESULT 12
 LOCUS AK056915 1839 bp mRNA linear PRI 01-AUG-2002
 DEFINITION Homo sapiens cDNA FLJ33535 fls, clone PROST2007612, highly similar
 to tricarboxylate carrier.
 ACCESSION AK056915
 VERSION AK056915.1 GI:16552444
 KEYWORDS oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1
 AUTHORS Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
 Ono,Y., Hotuta,T., Hirooka,S., Murakawa,K., Takiguchi,S.,
 Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
 Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
 Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y.,
 Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y.,
 Sekine,K., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fujii,A.,
 Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
 Nagahiri,K., Masuno,Y., Nagai,K. and Isogai,T.
 NEDO human cDNA sequencing project
 TITLE Unpublished
 JOURNAL (bases 1 to 1839)
 REFERENCE 2 (bases 1 to 1839)
 AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.
 TITLE Direct Submission
 JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@helix.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3985)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'-63'-end one pass sequenced by Japan
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: RAB and
 HRI.
 FEATURES
 source
 1. 1839
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PROST2007612"
 /issue_type="prostata"
 /clone_id="PROS12"
 /note="cloning vector: PME18SFL3"
 BASE COUNT 523 a 399 c 407 g 510 t

ORIGIN
 Alignment Scores:
 Pred. No.: 1,4e-91 Length: 1839
 Score: 1089.00 Matches: 201
 Percent Similarity: 89.45% Conservative: 28
 Best Local Similarity: 78.52% Mismatches: 27
 Query Match: 80.97% Indels: 0
 DB: 9 Gaps: 0
 US-09-990-415a-2 (1-261) x AK056915 (1-1839)
 QY 6 GlyIleuProLeuAspIleAsnIleGlnIleProArgTrpAspGlnSerPheLeu 25
 Db 63 GGAGAACATCCACCAACATTTACATCAAGAACCTCGATGGGATCAAGACCTTTCATT 122
 QY 26 GlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
 Db 123 GGACAGGCCAATCATTTCTTCACTGTAACAGCCACGAGAACATCTGTTAACCAACGAA 182
 QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
 Db 183 CAATCGAGAGTGCAGAAATAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 242
 QY 66 IleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisProAsp 85
 Db 243 CTTACAGAAATGATTTGTGAGACCAAGTACATCTTATGATTCAGCTTTTCATCTGAC 302
 QY 86 ThrGlyLysValValIleuIleGlyArgMetSerAlaGlnValProMetAsnMetThr 105
 Db 303 ACTGGTAGAAGATGATTTGTATGAGAAATGTCAGCCAGCTCCCATGACATGAC 362
 QY 106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125
 Db 363 ATCAGAGTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
 QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
 Db 423 ATTAAACAGCTCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482
 QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
 Db 483 ACTGTCATGAGTGTGGAAACACCTTACGTTTCTGCAACACCTGGCCCTACCAACAGCT 542
 QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
 Db 543 CTAGAGCATCAATGACATGACCAAGCATGCTCCACACGATGATGAGAGCTTTGTCCCTT 602
 QY 186 AlaAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 205
 Db 603 GCTGCCGTAAGTCTGCTGCTAATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 662
 QY 206 ValGlyIleProValAlaAlaAspGluAlaGlyIleArgLeuGlyTyrSerValThrAlaAla 225
 Db 663 GTTGGCATTTCCCGTCACGATGAGATGGAACCGCTTGGGGAGTCCGCGCAACCTCGC 722
 QY 226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
 Db 723 AACACAGCCATCACCAAGTGTGCTGTCAGAGATTCATGCGCAGCCCTGGCATGGCC 782
 QY 246 IleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeuLys 261
 Db 783 ATCCCTCATTCATTATGACACTTTGGAAGAAAGAGCCCTTTTGAAG 830
 RESULT 13
 LOCUS AX302019 2367 bp DNA linear PAT 30-NOV-2001
 DEFINITION Sequence 18 from Patent WO0177174.
 ACCESSION AX302019
 VERSION AX302019.1 GI:17382991
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens


```
Db      474  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Oy      146  ThrValargInleuGlyThrAlaTyValSerAlaThrThrAla 165
Db      534  AcTgAAATGAGCTGGAGACGCTTATGTTCCTGCACACCGCGCGTGGCAGACC 593
Oy      166  LeuGlyLeuLysSerLeuThrLysHisLeuProPheValGlyArgPheValProphe 185
Db      594  CTGGACTCATGGCCCTAACCAAGCATGTGCGCGCTGATAGACGTTTGTCCCTTC 653
Oy      186  AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetLArgGlnArgGln 205
Db      654  GCTGCTGAGCTGCTGCTAAGCTAATCCCGCTGATGAGACAAAGGAACTCAAG 713
Oy      206  ValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyrSerValThrAla 225
Db      714  GTTGGTATTCCTGTCACATGATGAGATGACACGCTTGGGAGATCGACCAACGACCA 773
Oy      226  LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMet 245
Db      774  AAGCAAGCCATCAGCAGAGTGTGTCATCTCCAGGATCCTCATGCGCCCTGGCATG 833
Oy      246  IleProPheLysIleMetAspThrLeuGlnLysLysAspPheLeuLys 261
Db      834  ATCCCTCCGTTTATCATGACACACCTTGAGAAAGAACCTTTTGAAG 881
```

Search completed: October 16, 2003, 16:59:43
Job time : 2923 secs

QY	349	GCTGCTGGCGCTCAACGGCTGACGTCGCCCGTGAATGGCGTGGAGAGGGCCCGGGGGGACACAGC	408
Db	146	GCTGCTGGCGCTCAACGGCTGACGTCGCCCGTGAATGGCGTGGAGAGGGCCCGGGGGGACACAGC	205
QY	409	GGAGGCAAGAGAGGAAGGCGGTTCTGAGAGCTTCCAGAGACATGGAAAGCAAAATGAGGTG	468
.Db	206	GGAGGCAAGAGAGGAAGGCGGTTCTGAGAGCTTCCAGAGAGCGATGGAAAGCAAAATGGGTG	265
QY	469	AATTGGCCTTTAGCAATCAACATCCAGAAACCTCGCTGGAGCCAAAGTACTTTCTGGGCA	528
Db	266	AATTGGCCTTTAGCAATCAACATCCAGAAACCTCGCTGGAGCCAAAGTACTTTCTGGGCA	325
QY	529	GAGCCCGGCACTTTTCTACTGTACTGATCCTCGAAATCTGCTGCTGTCGCGGGGCACAGC	588
Db	326	GAGCCCGGCACTTTTCTACTGTACTGATCCTCGAAATCTGCTGCTGTCGCGGGGCACAGC	385
QY	589	TGGAAGCTTCTCGGAACATCGTGCAGAACTACAGAGGCCGCGGTGGTGAAGCCACAGGATCA	648
Db	386	TGGAAGCTTCTCGGAACATCGTGCAGAACTACAGAGGCCGCGGTGGTGAAGCCACAGGATCA	445
QY	649	CCGAGAGCACACTGTGGAGGGCCAAAGTATGTATGATCTCCGCGCTTCCATCCGGACACAG	708
Db	446	CCGAGAGCACACTGTGGAGGGCCAAAGTATGTATGATCTCCGCGCTTCCATCCGGACACAG	505
QY	709	GGGAGAAAGTGTCCTGATTTGGCCGCATGTACGCCACAGGTCCCATGAACATGACATCA	768
Db	506	GGGAGAAAGTGTCCTGATTTGGCCGCATGTACGCCACAGGTCCCATGAACATGACATCA	565
QY	769	CTGGCTGCTATCTCACATTTCTACAGGAAGACCCCAACCGTGGTGTCTTGGCAGTGGGTGA	828
Db	566	CTGGCTGCTATCTCACATTTCTACAGGAAGACCCCAACCGTGGTGTCTTGGCAGTGGGTGA	625
QY	829	ATTCAGTCCCTTCAATGCCATTTGTAACTCAACTCCACCGAGTGGTGAACCTCCATCACTG	888
Db	626	ATTCAGTCCCTTCAATGCCATTTGTAACTCAACTCCACCGAGTGGTGAACCTCCATCACTG	685
QY	889	TGAGGCGAGCTGGGACACGCTATGTGAGTGGCCACCACTGAGACTGTGGGCCAGGCCCTGG	948
Db	686	TGAGGCGAGCTGGGACACGCTATGTGAGTGGCCACCACTGAGACTGTGGGCCAGGCCCTGG	745
QY	949	GACTCAAAATCCCTTACCAACAAGCACTGCCGCCCTTGGTGGGAGATTTGTGCCCTTTGGAG	1008
Db	746	GACTCAAAATCCCTTACCAACAAGCACTGCCGCCCTTGGTGGGAGATTTGTGCCCTTTGGAG	805
QY	1009	CAGTGGCAGCTGCCAATCTGATCAACATCCCCCTGATGAGGACAGAGAGCTGCAGGTGG	1068
Db	806	CAGTGGCAGCTGCCAATCTGATCAACATCCCCCTGATGAGGACAGAGAGCTGCAGGTGG	865
QY	1069	GCATCCCGGTTGGCGTATGAGGCAAGTGAAGGCTTGGTACTCGGTGACTGCAGCCAAAGC	1128
Db	866	GCATCCCGGTTGGCGTATGAGGCAAGTGAAGGCTTGGTACTCGGTGACTGCAGCCAAAGC	925
QY	1129	AGGGAATCTTCCAGAGTGGTGAATTTCAAGAAATCTCATGGCATTTCTCCATGGGCATCC	1188
Db	926	AGGGAATCTTCCAGAGTGGTGAATTTCAAGAAATCTCATGGCATTTCTCCATGGGCATCC	985
QY	1189	CACCACTGATATGACACATCTGGAGAAAGAAAGCTTCTGAAG	1232
Db	986	CACCACTGATATGACACATCTTGGAGAAAGAAAGCTTCTTGAAG	1029

US-08-905-223-165
 Sequence 165, Application US/08905223
 Patent No. 6222029
 GENERAL INFORMATION:
 APPLICANT: Edwards, Jean-Baptiste D.
 APPLICANT: Duclercq, Aymeric
 APPLICANT: Lacroix, Bruno
 TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
 NUMBER OF SEQUENCES: 503
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobbe, Martens, Olson & Bear

```

1 STREET: 501 West Broadway
2 CITY: San Diego
3 STATE: California
4 COUNTRY: USA
5 ZIP: 92101-3505
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Floppy Disk
9 COMPUTER: IBM PC compatible
10 OPERATING SYSTEM: Win95
11 SOFTWARE: Word
12
13 CURRENT APPLICATION DATA:
14 APPLICATION NUMBER: US/08/905,223
15 FILING DATE:
16 CLASSIFICATION: 536
17 ATTORNEY/AGENT INFORMATION:
18 NAME: Israelsen, Ned A.
19 REGISTRATION NUMBER: 29,655
20 REFERENCE/DOCKET NUMBER:
21 TELECOMMUNICATION INFORMATION:
22 TELEPHONE: (619) 235-8550
23 TELEFAX: (619) 235-0176
24 INFORMATION FOR SEQ ID NO: 165:
25
26 SEQUENCE CHARACTERISTICS:
27 LENGTH: 315 base pairs
28 TYPE: NUCLEIC ACID
29 STRANDEDNESS: DOUBLE
30 TOPOLOGY: LINEAR
31 MOLECULE TYPE: CDNA
32 ORIGINAL SOURCE:
33 ORGANISM: Homo Sapiens
34 TISSUE TYPE: Brain
35
36 FEATURE:
37 NAME/KEY: sig-peptide
38 LOCATION: 70..108
39 IDENTIFICATION METHOD: Von Heijne matrix
40 OTHER INFORMATION: score 3.5
41 OTHER INFORMATION: seq MHLSTMWNPASS/RR
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
9
```

```

RESULT 3
US-09-922-146-3
: Sequence 3, Application US/09922146
: Patent No. 6566133
: GENERAL INFORMATION:
: APPLICANT: Lex M. Cowser
: APPLICANT: Brett P. Motta
: TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 9 EXPRESSION
: FILE REFERENCE: RTS-0252
: CURRENT APPLICATION NUMBER: US/09/922,146
: CURRENT FILING DATE: 2001-08-01
: NUMBER OF SEQ ID NOS: 48
: SEQ ID NO 3

```

```

; LENGTH: 2303
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (114)..(1268)
US-09-922-146-3

```

Query Match	3.4%	Score 41.6;	DB 4;	Length 2303;
Best Local Similarity	51.3%;	Pred. No. 0.1;		
Matches 122;	Conservative 0;	Mismatches 114;	Indels 2;	Gaps 1.

QY 200 GCACCCCTTGAGAGCCCTCCGTGGAGGGAGACGTACAGGGGCAAGAGGGAGCGAGCTAGATT 259

Db 196 GCACCGCGGAGACTGTACGAGTCCGCGCCGATCGTGTGGGGCGCTGACCTGTGGCCCTTCCGG 255

QY 260 TGCCTGCAG--GCGTCCCTCTGTGTGGTGGCTGGGATTCTGCCAATCCCGCTGACCCACCGGTT 317

Db 256 CGCTCTCTGCTGCGCGCGCTGCGGAGAGGGGACGCTGTCGGTGGCGGCGGCTCTGCTGTGGG 315

QY 318 GGGGCGGGCCCGGGAAGCTCTGTCGCCCCCTCCTCTGCTGTGCGCGCTACGCGTACGTCGCC 377

Db 316 CGCGCGTCCACACCGCCGCCCGCTGCCCGCCCGTCTCTTATCACACCAAGGCGGGGCGCGCGC 375

QY 378 GTGATGGCTGGGAGGGCCCGCGGGGAGACAGCGGAGCACAAGAGAAAGGGGTTCTGAG 435

Db 376 GCCCGCGGGGGAGGCCAGGCGCCGAGCGCGAGAGTGGAGAGGCCGCTGTGCTCTGTGG 433

Query Match	3.3%	Score 41.2;	DB 3;	Length 1894;
Best Local Similarity	59.3%;	Pred. No. 0.12;		
Matches 70;	Conservative 0;	Mismatches 48;	Indels 0;	Gaps 0;

QY	305	GTCGCCACCGGGGGGGCGCGCCCGGGAAGTCTCTCCCTCCCTGTCGTGTGGCGGTACG	364
Db	14	GTGGGAAGAGGGTGGGGCGCAGAGGGGGAAGCCCCCGCGCCCTCTCCGGGCTTTGGGCGC	73
QY	365	CGGACAGTCCCGCGTGATGGCTGTGGAGAGGCCCGCGCGGACAGGGAGGACGAGAGA	422
Db	74	CGTGCCATTGGCCCCGGCGGCGCTTTGTGGGCGGGAAGATGACATCAGCGGCAAGTTGGA	131

; OTHER INFORMATION: "n" bases at various positions throughout the sequence
 ; OTHER INFORMATION: represent a, t, c or g
 US-09-103-840A-2

	Query Match	Similarity	3.3%	Score	0.8	DB	3	Length	4403765
	Best Local	Similarity	54.7%	Pred.	No.	3.7	Mismatches	81	Conservative
QY	302	CCCGTCGCCACCGGGTGGGCGCGCCGCGGAGACCTCTGCCCCCTGCTGCTGGCGGCTC	361						
Db	2210214	CCCGCGACACGACACACAGGGGGGTGCCAACACGAGGGGGAGCTCGAGTTTCGTTGTC	2210155						
QY	362	ACGCGTACAGCTCCCGCGTGATGCTGGGAGGGCCCCGGCGGCGACACGGAGGCAGAGAGG	421						
Db	2210154	ACGGTTGGTCTCCCGGTGCTGCCCGGCGGGGGCCGGCGGCGCGCCCGCGGCGGCGCGG	2210095						
QY	422	AAGCGGTCGAGAGCTTCGAGAGGG	449						
Db	2210094	CGGGCGGCTCGGCGCGGTACGAGATAGG	2210067						

```

? ATTORNEY/AGENT INFORMATION:
? NAME: SCOTT, WATSON T.
? REGISTRATION NUMBER: 26,581
? REFERENCE/DOCKET NUMBER: WTS/9437/91816
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202)861-3000
? TELEFAX: (202)822-0944
? TELEX: 671 4627 CUSH
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1894 base pairs
? TYPE: NUCLEIC ACID
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
US-07-728-220C-1

```

SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match
Best Local Similarity 54.7%; Score 40.8; DB 3; Length 4411529;
Matches 81; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

OY 302 CCCGTCACCCGCGGTGGGCGGCGGCGGGAAGCTCCGCTCCGCTGCTGGCGCTC 361
DB 2212914 CCCGCCACGACGACGAGCGGCGGCTCGCCACACGAGCGGCGACGTCTGAGTTGCTC 2212855
OY 362 ACGCGTAGCTCCCGCGGTGATGCTGGAGGCGCCGCGCGGCGAGCGAGCGAGAGG 421
DB 2212854 ACGGTGGTCTCCCGGCTGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2212795
OY 422 AAGCGGTCTGAGAGCTTACAGAGCG 449
DB 2212794 CGGCGGCTCCGCGGCGGCTAGCGATAGCG 2212767

RESULT 7
US-09-252-991A-7687/c
Sequence 7687, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7687
LENGTH: 642
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7687

Query Match
Best Local Similarity 48.5%; Score 39.8; DB 4; Length 642;
Matches 110; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

OY 779 GCTCACATCTACAGAAAGACCCCAACCGTGTCTTGGCAGTGGTAATCAGTCTT 838
DB 584 GCTGGCTTCGCCAAGAGAGTGGCGGCGGAGAAAGTCCACGCGCGGCGGCTGT 525
OY 839 CAATGCCATTGTAACTACTCAACCGCAGTGTGACACTCCATCAGTGTGAGCAGCT 898
DB 524 CCGTGGCATGGTGAAGAGCCCTACCGCAAGGCGCAAGGCGCTGAACATGGCCAGCTACT 465
OY 899 GGGGACAGCTATGTAGTGCACCACTGAGAGCTGTGGCCAGGCGCTGGAGCTCAATC 958
DB 464 GGAAGATCGAGCGGTATCGACCCGCGGAGACCCGCGCTGCTGCTGGCGGCGCTGGC 405
OY 959 CCGACCAAGCAGCCTGCCCCCTTGGTGGCAGATTTTGGCCCTTTG 1005
DB 404 CGTCGCCGAGAACCGCGCGCGGCGGCGGCGCAAGCGGCGCTTGC 358

RESULT 8
US-09-252-991A-7943
Sequence 7943, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7943
LENGTH: 1761
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7943

Query Match
Best Local Similarity 48.5%; Score 39.8; DB 4; Length 1761;
Matches 110; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

OY 779 GCTCACATCTACAGAAAGACCCCAACCGTGTCTTGGCAGTGGTAATCAGTCTT 838
DB 1360 GCTGGCTTCGCCAAGAGAGTGGCGGCGGAGAAAGTCCACGCGGCGGCGGCTGT 1419
OY 839 CAATGCCATTGTAACTACTCAACCGCAGTGTGACACTCCATCAGTGTGAGCAGCT 898
DB 1420 CCGTGGCATGGTGAAGAGCCCTACCGCAAGGCGCAAGGCGCTGAACATGGCCAGCTACT 1479
OY 899 GGGGACAGCTATGTAGTGCACCACTGAGAGCTGTGGCCAGGCGCTGGAGCTCAATC 958
DB 1480 GGAAGATCGAGCGGTATCGACCCGCGGAGAGACCCGCGCTGCTGCTGGCGGCGCTGGC 1539
OY 959 CCGTACCAGCAGCCTGCCCCCTTGGTGGCAGATTTTGGCCCTTTG 1005
DB 1540 CGTCGCCGAGAACCGCGCGGCGGCGGCGGCGCAAGCGGCGCTTGC 1586

RESULT 9
US-09-252-991A-7538/c
Sequence 7538, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7538
LENGTH: 2718
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7538

Query Match
Best Local Similarity 48.5%; Score 39.8; DB 4; Length 2718;
Matches 110; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

OY 779 GCTCACATCTACAGAAAGACCCCAACCGTGTGTCTTGGCAGTGGTAATCAGTCTT 838
DB 268 GCTGGCTTCGCCAAGAGAGTGGCGGCGGAGAAAGTCCACGCGGCGGCGGCTGT 209
OY 839 CAATGCCATTGTAACTACTCAACCGCAGTGTGACACTCCATCAGTGTGAGCAGCT 898
DB 208 CCGTGGCATGGTGAAGAGCCCTACCGCAAGGCGCAAGGCGCTGAACATGGCCAGCTACT 149
OY 899 GGGGACAGCTATGTAGTGCACCACTGAGAGCTGTGGCCAGGCGCTGGAGCTCAATC 958

Db 148 GGAGATGACGCGGTGATTCAGACCGGCGAGAGACCCGGCGCTGCTCGCGGCGCTGGC 89
QY 959 CCTCACCAGACACCTGCCCCCTTGTCTGGCGAGATTGTGCGCTTTG 1005
Db 88 CGTGGCGGAGAACCGCGCGCGAGGGGGGGCGCAAGCGGCGTTGC 42

RESULT 10

US-09-252-991A-7803
Sequence 7803, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7803
LENGTH: 3351
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7803

Query Match 3.2%; Score 39.8; DB 4; Length 3351;
Best Local Similarity 48.5%; Pred. No. 0.37; Indels 0; Gaps 0;

Matches 110; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 779 GCTCATTCTACAGAGAACCCCAACCGGTGTCTGGAGTGGGTGAATCAAGTCTT 838
Db 3111 GCTGGGCTTCCGACAGAGCTGGGGCGCGAGAGAGTCCCAAGCGCGCGCTTT 3170
QY 839 CAATGCAATTTTAACTACTCAACCGCATGTGTGACACTCCACTGTGAGGCACT 898
Db 3171 CCGTGGCATGTGAGCAAGGCTTACCCAGCAAGCGCGCTGAACATGGCCAGCTACT 3230
QY 899 GGGAGACGCTATGTGATGTCACACCTGAGAGCTGTGGCCAGCGCCCTGGGATCAATC 958
Db 3231 GGAATGACCGCGGTGATGACCGCGCGAGAACCCCGCGTGGCTCTCGCGCGGCTTGC 3290
QY 959 CCTCACCAGACACCTGCCCCCTTGTCTGGCGAGATTGTGCGCTTTG 1005
Db 3291 CGTGGCGGAGAACCGCGCGCGAGGGGGGGCGCAAGCGGCGCTTGC 3337

RESULT 11

US-08-306-691B-14
Sequence 14, Application US/08306691B
Patent No. 5734039
GENERAL INFORMATION:
APPLICANT: Calabretta, Bruno
APPLICANT: Skorski, Tomasz
TITLE OF INVENTION: ANTISENSE
TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seidel, Gonda, Lavorina & Monaco, P. C.
STREET: Two Penn Center, Suite 1800
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/306,691B
FILING DATE: September 15, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5734039e

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
LENGTH: 6453 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

Query Match 3.1%; Score 38.4; DB 1; Length 6453;
Best Local Similarity 44.8%; Pred. No. 1.2; Indels 0; Gaps 0;

Matches 147; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 87 GCTCATTCCGCAAAATTGGGCTTGGGCTGTGGCTGTTCCTCCCTCGGCTCGAACCACCTC 146
Db 576 GGTGAGTCCGGGACAGAGATCGCGCGCGCGCGCGCGCGCTCTCGCCCGACGGCGAGCAA 635
QY 147 TTCTGTGAGCCGAGCAGCTACCGGGGCTCTGTGAATTCACACCCCTCTGGGACACT 206
Db 636 TACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTAAGCGCGCGCGCGCGCG 695
QY 207 TGAGGCTCCGTTGAGGAGGAGACGTACAGGGGAGCGGAGCGAGCTGAGTTGCTGCA 266
Db 696 GGTGGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCGGGGCTCGGCG 755
QY 267 GCGGTGCTGT 326
Db 756 GCGCGGGCG 815
QY 327 CGGAGAGCTCTGCTCCCTTCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 386
Db 816 CTGGGACAGGTCCTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 875
QY 387 GGAAGGCG 414
Db 876 GGGCGCTGTGGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 903

RESULT 12

US-09-209-668-10
Sequence 10, Application US/09209668A
Patent No. 6114517
GENERAL INFORMATION:
APPLICANT: Monia, Brett P.
APPLICANT: Xu, Xiaoxing S.
TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
TITLE OF INVENTION: ALPHA-INDUCED EXPRESSION OF CELL ADHESION MOLECULES
FILE REFERENCE: ISPH-0336
CURRENT APPLICATION NUMBER: US/09/209,668A
CURRENT FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 6453
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1664)..(1774)
FEATURE:

TITLE Human transport proteins
JOURNAL Patent: WO 0078953-A 46 28-Dec-2000;
Incyte Genomics, Inc. (US)
FEATURES Location/Qualifiers
source 1..1826
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 415 a 519 c 512 g 380 t
ORIGIN

Query Match 99.9%; Score 1231; DB 6; Length 1826;
Best Local Similarity 100.0%; Pred. No. 7.6e-268;
Matches 1231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCTTAGGCGCCAGGAGACAGCCGATTACCTGGTCCCGGACAGCGAGTTCTTTACCA 61
Db 1 CCTTAGGCGCCAGGAGACAGCCGATTACCTGGTCCCGGACAGCGAGTTCTTTACCA 60
QY 62 CCCAGATTCTGTGTGACGCCCTTACGCTCATTCGCCAATTAGGCGTTGGGCTGGCTT 121
Db 61 CCCAGATTCTGTGTGACGCCCTTACGCTCATTCGCCAATTAGGCGTTGGGCTGGCTT 120
QY 122 GTTCCCTCCGCGTGAACACACCTTCTGTAGCCGAGCAGCTACCGGGGCTCCTGA 181
Db 121 GTTCCCTCCGCGTGAACACACCTTCTGTAGCCGAGCAGCTACCGGGGCTCCTGA 180
QY 182 ATTGCCACCCCTCCCTGGGACACCTTGAGGCTCCGTTGAGGAGCAGTACGGGGCAGAGC 241
Db 181 ATTGCCACCCCTCCCTGGGACACCTTGAGGCTCCGTTGAGGAGCAGTACGGGGCAGAGC 240
QY 242 GGGAGGAGGCTGAGTTGTGTGAGGCTGCTGTGTGGTGGTGGTCTGCTCCATC 301
Db 241 GGGAGGAGGCTGAGTTGTGTGAGGCTGCTGTGTGGTGGTGGTCTGCTCCATC 300
QY 302 CCCGTGCCACCGGGGTGGGCGCGGCGGGAAGCTCTGCCCTCCCTGCTGTGGGCTGC 361
Db 301 CCCGTGCCACCGGGGTGGGCGCGGCGGGAAGCTCTGCCCTCCCTGCTGTGGGCTGC 360
QY 362 ACCGCTGACGTCGCCGCTGATGGCTGGAGGGCCCGGCGGACAGCCGAGGACAGAGG 421
Db 361 ACCGCTGACGTCGCCGCTGATGGCTGGAGGGCCCGGCGGACAGCCGAGGACAGAGG 420
QY 422 AAGCGGTTCTGAGAGCTTCAAGAGAGCATGAAGAAATGGGTGAATGGCTTTGA 481
Db 421 AAGCGGTTCTGAGAGCTTCAAGAGAGCATGAAGAAATGGGTGAATGGCTTTGA 480
QY 482 CATCAACATCCAGGAACCTGCTGGGACCAAGTACTTCTGGGACAGGCCGCGCACTT 541
Db 481 CATCAACATCCAGGAACCTGCTGGGACCAAGTACTTCTGGGACAGGCCGCGCACTT 540
QY 542 TTTGACGTTACTGATCTCGAAATCTGCTGTCGCGGGGACAGCTGGAAGCTTTCG 601
Db 541 TTTGACGTTACTGATCTCGAAATCTGCTGTCGCGGGGACAGCTGGAAGCTTTCG 600
QY 602 GAACATCTGCAAGAACTACAGGGCGGGGCTGTGATGCCAGGATCACCGAGACACT 661
Db 601 GAACATCTGCAAGAACTACAGGGCGGGGCTGTGATGCCAGGATCACCGAGACACT 660
QY 662 GTGAGGGCCAAAGTATGTATGACTCCGCTTCATCCGAGACAGGGGGAAGTGT 721
Db 661 GTGAGGGCCAAAGTATGTATGACTCCGCTTCATCCGAGACAGGGGGAAGTGT 720
QY 722 CCTGATTGGCGGCAATGTCAGCCAGGTGCCATGAACATGACACATCGCTGTGATCT 781
Db 721 CCTGATTGGCGGCAATGTCAGCCAGGTGCCATGAACATGACACATCGCTGTGATCT 780
QY 782 CACATTTACAGAGAACCCCAACCGTGTGTTCTGGCAGTGGGTAATCAATCTTAA 841
Db 781 CACATTTACAGAGAACCCCAACCGTGTGTTCTGGCAGTGGGTAATCAATCTTAA 840
QY 842 TGCCATTGTAACTACTCCAAACCGAGTGTGACACTCCATCTGTGAGGACAGCTGG 901
Db 841 TGCCATTGTAACTACTCCAAACCGAGTGTGACACTCCATCTGTGAGGACAGCTGG 900

Db 841 TGCCATTGTAACTACTCCAAACCGAGTGTGACACTCCATCTGTGAGGACAGCTGG 900
QY 902 GACAGCCTATGTGATGCGCACCACTGAGCTGTGGCCAGGCGCTGGGACTCAATCCCT 961
Db 901 GACAGCCTATGTGATGCGCACCACTGAGCTGTGGCCAGGCGCTGGGACTCAATCCCT 960
QY 962 CACCAAGCACTGCCCGCTTGTGTCGAGATTTGTGCTTTTGACAGATGGCAGCTGC 1021
Db 961 CACCAAGCACTGCCCGCTTGTGTCGAGATTTGTGCTTTTGACAGATGGCAGCTGC 1020
QY 1022 CAATGCATCAACATCCCGCTGATGAGGACAGAGAGCTGAGGTGGGCTCCGCTGGC 1081
Db 1021 CAATGCATCAACATCCCGCTGATGAGGACAGAGAGCTGAGGTGGGCTCCGCTGGC 1080
QY 1082 TGATGAGGACAGTCAAGAGCTTGGCTACTCGGTGACTGACAGCAAGAGGAATCTTCA 1141
Db 1081 TGATGAGGACAGTCAAGAGCTTGGCTACTCGGTGACTGACAGCAAGAGGAATCTTCA 1140
QY 1142 GGTGTGATTTCAAGAAATCTGCATGGCATTCCTGCCATGGCCATCCACATGATCAT 1201
Db 1141 GGTGTGATTTCAAGAAATCTGCATGGCATTCCTGCCATGGCCATCCACATGATCAT 1200
QY 1202 GGACACTCTGGAGAAAGAAAGACTTCTGAAG 1232
Db 1201 GGACACTCTGGAGAAAGAAAGACTTCTGAAG 1231

RESULT 2

AX188304/c AX188304 3047 bp. DNA linear PAT 06-AUG-2001

LOCUS Sequence 3999 from Patent WO0142467.

DEFINITION AX188304

ACCESSION AX188304.1 GI:15139777

VERSION AX188304.1

KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE 1

AUTHORS Schlegel, R., Deeds, J., Berger, A. and Zhao, X.

TITLE Genes, compositions, kits, and methods for identification,

JOURNAL assessment, prevention, and therapy of cervical cancer

Patent: WO 0142467-A 3999 14-JUN-2001;

Millennium Predictive Medicine, Inc. (US)

FEATURES Location/Qualifiers

source 1..3047

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 629 a 790 c 891 g 733 t 4 others

ORIGIN

Query Match 98.1%; Score 1208.2; DB 6; Length 3047;
Best Local Similarity 99.8%; Pred. No. 1.1e-262;
Matches 1210; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 20 GCCAGCGTTACGCTGCTCCGCGGACGCGAGTCTTACCAACCCAGTTGGTTGCA 79
Db 3035 GTCCGCGCTTACGCTGCTCCGCGGACGCGAGTCTTACCAACCCAGTTGGTTGCA 2976
QY 80 CGCCCTAGCATATTCGCAAAATTTAGGGCTTGGTCTGTGTTCCCTCGGCTCGAA 139
Db 2975 CGCCCTAGCATATTCGCAAAATTTAGGGCTTGGTCTGTGTTCCCTCGGCTCGAA 2916
QY 140 CCACCTCTTCTGTAGCGCGAGCAGCTACCGGGCTCTCGAATTTGCCACCCCTCCCTGG 199
Db 2915 CCACCTCTTCTGTAGCGCGAGCAGCTACCGGGCTCTCGAATTTGCCACCCCTCCCTGG 2856
QY 200 GCACCTTGAGGCGCTCGTGGAGGAGCTACAGGGGAGAGGCGGAGAGCTGAGCTGAGTT 259
Db 2855 GCACCTTGAGGCGCTCGTGGAGGAGCTACAGGGGAGAGGCGGAGAGCTGAGCTGAGTT 2796
QY 260 TGCTGACAGCGTGTCTGTGTGTGCTGTGGGTTCTGCACAAATCCCGTCCACCGGGTGG 319

LOCUS	BD127742	2702 bp	DNA	linear	PAT 18-SEP-2002
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.				
ACCESSION	BD127742				
Db	2795	TGTCGACAGGCGCTGTGTGTGTGGGCTTCTGCCAATCCCGTGTCCACCGGGGTGG	2736		
OY	320	GGCGGGCCGGGAAAGCTCTGTGCCCTCCTGCTGGTCCGCGTCACGCGGTACGTCGCCGT	379		
Db	2735	GGCGGGCCGGGAAAGCTCTGTGCCCTCCTCCTCCTGTGGTCGCGCTCACGCGTGAAGTCCCGGT	2676		
OY	380	GATGGCTGGAGGGGCCCGGCGGACAGACGGAGGAGAGAGAAAGCGGTTCTGAGAGCT	439		
Db	2675	GATGGCTGGAGGGGCCCGGCGGACAGACGGAGGAGAGAGAAAGCGGTTCTGAGAGCT	2616		
OY	440	TCAGAGACCGATGGAAAGCAAAATGGGTGAATTGCCTTTAGACATCAACATCCAGAAC	499		
Db	2615	TCAGAGACCGATGGAAAGCAAAATGGGTGAATTGCCTTTAGACATCAACATCCAGAAC	2556		
OY	500	TCGCTGGGACCAAAAGTACTTCTCTGGGAGAGCCCGGACACTTTTTCACCTGTACTGATCC	559		
Db	2555	TCGCTGGGACCAAAAGTACTTCTCTGGGAGAGCCCGGACACTTTTTCACCTGTACTGATCC	2496		
OY	560	TCGAATCTGCTGCTGTCCGGGGACAGACTGGAAGCTTCGGAACATCGACAGACTA	619		
Db	2495	TCGAATCTGCTGCTGTCCGGGGACAGACTGGAAGCTTCGGAACATCGACAGACTA	2436		
OY	620	CAGGGCCGGCGCTGTGTACCCAGGATCCAGAGACCAAGCTGTGGAGGCCAATGAT	679		
Db	2435	CAGGGCCGGCGCTGTGTACCCAGGATCCAGAGACCAAGCTGTGGAGGCCAATGAT	2316		
OY	680	GTATGACATCCGCTCCATCCGGGACACAGGGGAGAAAGGTGGTCTGATATGGCCCATGTC	739		
Db	2375	GTATGACATCCGCTCCATCCGGGACACAGGGGAGAAAGGTGGTCTGATATGGCCCATGTC	2316		
OY	740	AGCCAGGTGCCATGAAACATGACCATCTACTGCTGATGCTCATCTTCATTCACAGAAAGAC	799		
Db	2315	AGCCAGGTGCCATGAAACATGACCATCTACTGCTGATGCTCATCTTCATTCACAGAAAGAC	2256		
OY	800	CCCAACCGTGGTGTTCCTGGCAGTGGGTGAATCACTCTTCAATCCATGTTAACTACTC	859		
Db	2255	CCCAACCGTGGTGTTCCTGGCAGTGGGTGAATCACTCTTCAATGTTAACTACTC	2196		
OY	860	CAACCGCAGTGGTGAACACTCCCATCTGTTGAGCAGCTGGGAGACGCTATGTGAGTGC	919		
Db	2195	CAACCGCAGTGGTGAACACTCCCATCTGTTGAGCAGCTGGGAGACGCTATGTGAGTGC	2136		
OY	920	CACCACTGGAGCTGTGGCCACAGGCGCTGGGACTAAATCCCTCACCAAGCACTGCCCC	979		
Db	2135	CACCACTGGAGCTGTGGCCACAGGCGCTGGGACTAAATCCCTCACCAAGCACTGCCCC	2076		
OY	980	CTTGGTGGGCAATTTGTGCCCTTTGGAGCAGTGGGACGTCGCAACGCAATCAACATCCC	1039		
Db	2075	CTTGGTGGGCAATTTGTGCCCTTTGGAGCAGTGGGACGTCGCAACGCAATCAACATCCC	2016		
OY	1040	CCTGATGAGGACAGAGAGAGCTGACAGTGGGCAATCCGGTGGCTGATGAGGACGTCAGAG	1099		
Db	2015	CCTGATGAGGACAGAGAGAGCTGACAGTGGGCAATCCGGTGGCTGATGAGGACGTCAGAG	1956		
OY	1100	GCTTGGCTACTCCGTTGACTGACAGCAAGCAGAGGAATCTTCCAGTGGTGAATTCAGAAAT	1159		
Db	1955	GCTTGGCTACTCCGTTGACTGACAGCAAGCAGAGGAATCTTCCAGTGGTGAATTCAGAAAT	1896		
OY	1160	CTGCGATGGCAATTCCTGCAATGGGCAATCCACCACTGATCATGGACAACCTGGAGAAAG	1219		
Db	1895	CTGCGATGGCAATTCCTGCAATGGGCAATCCACCACTGATCATGGACAACCTGGAGAAAG	1836		
OY	1220	AGACTTCTCTGAAG 1232			
Db	1835	AGACTTCTCTGAAG 1823			

VERSION	BDJ27742.1	GI:23222687
KEYWORDS	JP 2002017375-A/3173.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.	
TITLE	Primer for synthesizing full-length cDNA and use thereof	
JOURNAL	Patent: JP 2002017375-A 3173 22-JAN-2002;	
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002017375-A/3173 PD 22-JAN-2002 PF 07-JUL-2000 JP 2000253172 PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO PI ISHII, PI YURI KAWAI,AI WAKAMATSU,TOMOTASU SUGIYAMA,KEIICHI MAGAI, PI SHINICHI KOJIMA, PI TETSUJI OTSUKI,HISASHI KOGA PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10, C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers FT CDS Location/Qualifiers 1..2702 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"	
BASE COUNT	608 a 757 c 699 g 638 t	
ORIGIN		
Query Match	68.9%; Score 848.8; DB 6; Length 2702;	
Best Local Similarity	99.8%; Pred. No. 2.1e-181;	
Matches	850; Conservative 0; Mismatches 2; Indels 0; Gaps 0	
Dp	1 ATGGCTGGAGAGGCCCGGCGGCGAGAGCGAGAGAGAGAAAGCGTTCTGAGAAGTTT 440	
Qy	381 ATGGCTGGAGAGGCCCGGCGGCGAGAGCGAGAGAGAGAGAAAGCGTTCTGAGAAGTTT 440	
Dp	1 ATGGCTGGAGAGGCCCGGCGGCGAGAGCGAGAGAGAGAGAAAGCGTTCTGAGAAGTTT 60	
Qy	441 CAGAGAGCGATGGAAGCAAAATGGGTGAATTCCCTTTAGACATCAACATCCAGAACCT 500	
Dp	61 CAGAGAGCGATGGAAGCAAAATGGGTGAATTCCCTTTAGACATCAACATCCAGAACCT 120	
Qy	501 CGCTGGGAGCAAAGACTTTCCTGGGGCAGAGCGCGGCACTTTTCACTGTTACTGACT 560	
Dp	121 CGCTGGGAGCAAAGACTTTCCTGGGGCAGAGCGCGGCACTTTTCACTGTTACTGACT 180	
Qy	561 CGAATCTGCTGCTCTCCGGGGGCACAGCTGGAAGCTTTCGGAACATGTCGAGAACTAC 620	
Dp	181 CGAATCTGCTGCTCTCCGGGGGCACAGCTGGAAGCTTTCGGAACATGTCGAGAACTAC 240	
Qy	621 AGGGCCGGCGTGCTGACCCCAAGGGATCACCCAGAGACAAGCTGTGAAGGGCCAATATGTG 680	
Dp	241 AGGGCCGGCGTGCTGACCCCAAGGGATCACCCAGAGACAAGCTGTGAAGGGCCAATATGTG 300	
Qy	681 TATGACATCCGCTTCATCGGAGACAGAGGGGGAAGGTGCTGATTTGGCCGCATGTCA 740	
Dp	301 TATGACATCCGCTTCATCGGAGACAGAGGGGGAAGGTGCTGATTTGGCCGCATGTCA 360	
Qy	741 GCCCAGGTGCCCATGAACATGACCATCTGAGTGCATGCTTCACTTTCAGAGGAAGACC 800	
Dp	361 GCCCAGGTGCCCATGAACATGACCATCTGAGTGCATGCTTCACTTTCAGAGGAAGACC 420	
Qy	801 CCAACCGGTGGTCTCGGCAATGGGTGAATAAGTCCCTTCAATGCATGTTTAACATATCC 860	
Dp	421 CCAACCGGTGGTCTCGGCAATGGGTGAATAAGTCCCTTCAATGCATGTTTAACATATCC 480	

QY	861	AAACGCGATGGTGCACACACCCCATCATCACTGTGAGGAGCGTGGGAGACAGCCTATGTAGTGGC	920
Db	481	AAACGCGATGGTGCACACACCCCATCATCACTGTGAGGAGCGTGGGAGACAGCCTATGTAGTGGC	540
QY	921	ACACATGGAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTACACAGACACCTGCCCCC	980
Db	541	ACCACTGGAGCTGTGGCCACGGCCCTGGGACTCAAAATCCCTACACAGACACCTGCCCCC	600
QY	981	TTGGTCGGCACAATTGTGGCCCTTTGGCGACACTGGCAGCTGGCAACTGCATTAACATCCCC	1040
Db	601	TTGGTCGGCACAATTGTGGCCCTTTGGCGACACTGGCAGCTGGCAACTGCATTAACATCCCC	660
QY	1041	CTGATGAGGCGAGAGAGAGCTGCAGGTGGGATCCCGGGTGGCTGATGAGAGCAGTCAAGG	1100
Db	661	CTGATGAGGCGAGAGAGAGCTGCAGGTGGGATCCCGGGTGGCTGATGAGAGCAGTCAAGG	720
QY	1101	CTTGGGCTACTGTGGTGATCGACAGCCACAGGGAATCTTCCAGGTGGTGATTTCAAGAATC	1160
Db	721	CTTGGGCTACTGTGGTGATCGACAGCCACAGGGAATCTTCCAGGTGGTGATTTCAAGAATC	780
QY	1161	TGCATGGCGGATTCCTGGCCATGGCCATCCACACACTGATCATGAGACACTGTGAAGAA	1220
Db	781	TGCATGGCGGATTCCTGGCCATGGCCATCCACACACTGATCATGAGACACTGTGAAGAA	840
QY	1221	GACCTTCCTGGAG 1232	
Db	841	GACCTTCCTGGAG 852	
RESULT 4			
LOCUS	AK074707	2702 bp	mRNA linear PRI 03-SEP-2002
DEFINITION	Homo sapiens cDNA FLJ90226 fls, clone NT2RM1000858, weakly similar to tricarboxylate carrier.		
ACCESSION	AK074707		
VERSION	AK074707.1	GI:22760323	
KEYWORDS	oligo capping; fls (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T., Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S., Kawai,Hio,Y., Saito,K., Yamamoto,J., Makamatsu,A., Nakamura,Y., Kojima,S., Nagahari,K., Masuno,Y., Ono,T., Okuno,K., Yoshikawa,Y., Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and Niimura,K.		
TITLE	NEDO human cDNA sequencing project		
JOURNAL	unpublished		
REFERENCE	2 (bases 1 to 2702)		
AUTHORS	Isogai,T. and Otsuki,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hrl.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)		
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).		
FEATURES	Location/Qualifiers		
source	1..2702		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="NT2RM1000858"		
	/cell_line="NT2"		
	/cell_type="teratocarcinoma"		
	/clone_id="NT2RM1"		
	/note="Cloning vector: pUC19FLJ3-mRNA from uninduced NT2"		

CDS		neuronal precursor cells"
	70..1047	/note="unnamed protein product"
		/codon_start=1
		/protein_id="BAC1151.1"
		/db_xref="gi:22760324"
		/translation="MESMGELPLDINIOEPRMDSTFLGRHFETVTPDRNLISG AQLSEBNINYRAGYVRIGTIEDOLMRAYKSDSPHPDETCKYVLIRMSQVPM NMITTCGLNTRKTPTTVYERFVNQSNALVYN SNSNGDTPIVRLGTRYSATLG AVATALGIKSLTKHLPLVGERFEPFAVAANLCINTLMQRLEQVI PVADENGQEL GYSTAAKGIGIYIVLISRICAIDPAMAIPIMIDPLEIKHCFKLKRPMGLPVLTGVG FCLVATPLPCALLEPKOKSHISMSELRAQIHEDNPSEVVYYNKGIL"
BASE COUNT	608 a 757 c 699 g 638 t	
ORIGIN		
Query Match	68.9%;	Score 848.8; DB 9; Length 2702;
Best Local Similarity	99.8%;	Pred. No. 2.1e-181;
Matches	Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Dy	381 ATGGCTGGAGAGGCCCGGCGCGCAGACAGCGGAGAGAGAGAAGCGGTCTCGAGACTT	440
Dd	1 ATGGCTGGAGAGGCCCGGCGCGCAGACAGCGGAGAGAGAAGCGGTCTCGAGACTT	60
Oy	441 CAGAGAGCATGTGAAGCAAAAATGGGTGAATTGCCTTTAGACATCAACATCCAGAACT	500
Dd	61 CAGAGAGCATGTGAAGCAAAAATGGGTGAATTGCCTTTAGACATCAACATCCAGAACT	120
Oy	501 CGCTGGGACCAAATAGCTTCTTGAGGACAGACCAGCACTTTTCACTGTACTGATCT	560
Dd	121 CGCTGGGACCAAATAGCTTCTTGAGGACAGACCAGCACTTTTCACTGTACTGATCT	180
Oy	561 CGAAATCTGCTGTGTCGCGGGGCGACAGCTGGAACTTCTCGGAACATCTGTCAGACTAC	620
Dd	181 CGAAATCTGCTGTGTCGCGGGGCGACAGCTGGAACTTCTCGGAACATCTGTCAGACTAC	240
Oy	621 AGGGCCGGCGGTGGAGACCCCGAGGATATCACGAGAACAGCTGTGAGAGGCCAAGTAGTG	680
Dd	241 AGGGCCGGCGGTGGAGACCCCGAGGATATCACGAGAACAGCTGTGAGAGGCCAAGTAGTG	300
Oy	681 TATGACTCCGCTTCCATCCGGACACAGAGGAGAAGGTGTCCTGATTGGCCGATGTCA	740
Dd	301 TATGACTCCGCTTCCATCCGGACACAGAGGAGAAGGTGTCCTGATTGGCCGATGTCA	360
Oy	741 GCCCAGGTGCCATGAACATGACCATACACACGCTGCACATGCTCAATCTACAGAGACC	800
Dd	361 GCCCAGGTGCCATGAACATGACCATACACACGCTGCACATGCTCAATCTACAGAGACC	420
Oy	801 CCAACCGTGTGTCTCTGGCAGTGGTGAATCAGTCTTTCATAGCATGTGTAATACTGCC	860
Dd	421 CCAACCGTGTGTCTCTGGCAGTGGTGAATCAGTCTTTCATAGCATGTGTAATACTGCC	480
Oy	861 AACCGCAGTGTGACAATCCCATATCAGTGTGAGCAGCTGGGAGACGCTATGTAGTCCC	920
Dd	481 AACCGCAGTGTGACAATCCCATATCAGTGTGAGCAGCTGGGAGACGCTATGTAGTCCC	540
Oy	921 ACCACTGAGAGTGTGGCCACGCGCCCTGGAGTCAAACTCCACACAGACACTGCCCCC	980
Dd	541 ACCACTGAGAGTGTGGCCACGCGCCCTGGAGTCAAACTCCCTCACACAGACTGCCCCC	600
Oy	981 TTGGTGGCAGATTGTGCCCCCTTGGACAGCAGTGGCAGCTGCCAACTGCATCAACATCCC	1040
Dd	601 TTGGTGGCAGATTGTGCCCCCTTGGACAGCAGTGGCAGCTGCCAACTGCATCAACATCCC	660
Oy	1041 CTGATGAGGAGAGAGAGAGCTGACAGGTGGGATCCCGGTGGCTGATGAGGAGAGCTCAGG	1100
Dd	661 CTGATGAGGAGAGAGAGAGCTGACAGGTGGGATCCCGGTGGCTGATGAGGAGAGCTCAGG	720
Oy	1101 CTGGCTACTCTCGGTAGCTGACGCCAACACAGGGAATCTTCCACAGTGTGATTTCAGATC	1160
Dd	721 CTGGCTACTCTCGGTAGCTGACGCCAACACAGGGAATCTTCCACAGTGTGATTTCAGATC	780
Oy	1161 TGCAATGGCGATTCTCGGCATGGCCATCCACACACTATCATGGACACTCTGGAGAGAAA	1220


```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="MGC:18958 IMAGE:3985155"
/tissue_type="Mammary tumor. WAP-RGF alpha model. 7 months
old, gross tissue."
/clone_1lb="NCI_CGAP_Mam5"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2870
/gene="Sfxn3"
/db_xref="LocusID:94280"
/db_xref="MGI:2137679"
221. .1186
/codon_start=1
/product="sideroflexin 3"
/protein_id="AAH12208.1"
/db_xref="GI:15126554"
/db_xref="LocusID:94280"
/translaton="MSDPLPLNTINIOEPRMDSFELGRARHFTVTDPERNLISGEOLE
ASRNIVONYRAGVATPGLETDLRAKYYDSAFHPDTGEKVVLIIGMSASVPANMTI
TSCMLFYRKPTFYVFMQWNOSFNALIVNTSNRSGADPTVVOGLGTASVATGAVAT
ALGKSLTKHLPLVGRFVPPAAVAAANCINIPIMRDELQNGCPYDDESGORGHGV
TAAGGIFGVYISRGIAIPAKAIPPIYNNITLEKDDFKRPFMIGAPLQVLGVFCFLV
FAIPPLCCALPQSSIHVTRLEPRLRAQIOAQNPSIDVYVYNGKL"

```

Query Match	Similarity	Score	DB ID	Length
Best Local	85.4%	Precl. No. 7.2e-142		
Matches	786	Conservative	0	Mismatches 130; Indels 4; Gaps 3;
QY	313	CGGGTGGGCGCGCGGGGAAAGCTCTCCCTCCCTCTGGTGGCGTCAACGGGTAGCT	372	
Db	76	CAGCGCAACGACGCTCAGGAGACCCCGCCCTTCCCTCTCTCAGCGTCAACGGGTAGCT	135	
QY	373	CCGCGGTGATGGCTGGAGGCGCCGCGCGACAGCGAGGACAGAGGAGGAGCGGTTCT	432	
Db	136	CTCG-CTGTAATGGCTGGAGGAGAAA--GCGGAGAGCGGTGAGGAGGAGCGGCTTGAGAG-CT	191	
QY	433	GAGAGCTTCAAGAGGCGATGGAAAGAAATGGGTGAATGGCTTTAGACATCAACATCC	492	
Db	192	TCTTAGAGGCTGAAAGACCCCGAAAGCAAGATGGGTGACGTCCCTTAATATCAACATCC	251	
QY	493	AGGAACCTGCTGGGACCAAGTACTTCTCTGGGACAGAGCCCGGCACTTTTCACTGTTA	552	
Db	252	AGGAACCTCGGTGGGACCAAGCAATTTTCAAGCAAGCCCGGCACTTTTCTCAGCTCA	311	
QY	553	CTGATCTCGAAATCTCTCTGCTGTCGGGGACAGCTGGAAGCTTTCGGAACATCGTGC	612	
Db	312	CTGATCCCGCAAACTCTCTCTGCTGTCGGGGACAGCTGGAAGCTTTCGGAACATCGTGC	371	
QY	613	AGAACTACAGGGCGGGGTGGTGCACCCACAGGAACTACCGAGAGACACTCTGAGAGGCCA	672	
Db	372	AGAACTACAGGGCGGTGGTGGTGCACCCCGGTCTCATCTGAGAGACACTTATGGGAGGCCA	431	
QY	673	AGATATGTATGATCCGCTTCCATCCGAGACAGGGGAGAAAGTGTCTCATATTGGCC	732	
Db	432	AATACGTATGATCACTAGCATTCATCCATCCGAGACAGGGGAGAAAGTGTCTTATATTGGCC	491	
QY	733	GCATGTACGGCCAGAGTCCCATTAACATGACATCACTGGCTGATGCTCTACATTTCTACA	792	
Db	492	GTATGTAGGCGCAGGTGCCCATGAACATGACATCACTTGGCTGATGCTCTACATTTCTACA	551	
QY	793	GGAAGACCCCAACCGTGGTTCGAGGAGGAGGGAATCAAGTCCCTCAATGCCATTGTTA	852	
Db	552	GGAAGACTCGAGCTGTGTTCGAGGAGGAGGGAATCAAGTCCCTCAATGCTATTGTTGA	611	
QY	853	ACTACTCAACCGCAGTGTGACACTCCATCACTGTGAGGACAGCTGGGAGACGCTATG	912	
Db	612	ATTACTTAATTCGAGGCGGATGCTCCATCACTGTGACAGCAAGTGGGAGACGCTATG	671	
QY	913	TGATGTCACACGCTGGAGCTGTGGCCACGGGCCCTGGGATCAAAATCCCTACCAAGACC	972	

Df	672	TGAGTGGCCACCACTGGGGGTGTGGCTACTGGCTTGCGGACTGAAGTCCTCACAAGAACC	731		
QY	973	TGCCCCCTTGGTCGGCAGATTTTGTGCCCTTTGCAGCAGTGGACCTGCCAATGTGATCA	1032		
Df	732	TGCCCCCCCTGTGGTGTGATTCGGCCCTTTCAGCTGAGCTGGCCCTCCAACTGCATCA	791		
QY	1033	ACATCCCCCTGTATGAGCAGAGAGAGTGCAGGTGGAGCATCCGGGTGGCTGTATGAGGAG	1092		
Df	792	ACATCCCCCTGTATGAGCAGAGAGAGTGCAGGTGGAGCATCCGGGTGGCTGTATGAGGAG	851		
QY	1093	GTCAGAGCTTGGCTACTCGGTGATCTGCATGCAAGCCAAAGCAGGAGATCTTCAGGTGTATTT	1152		
Df	852	GTCAGAGCTTGGCCACTCGGTGATCTGTCCCAAACAGAGGATCTTCAGGTGTATTT	911		
QY	1153	CAGAAATGTGCATGGCGAATTCCTGCAATGGCCATCCCACTATATATGACACTCTGG	1212		
Df	912	CAGAAATGTGCATGGCGAATTCCTGCAATGGCCATCCCACTATATATGACACTCTGG	971		
QY	1213	AGAGAAAGACTTCTCTGAAG	1232		
Df	972	AGAGAAAGACTTCTCTGAAG	991		
RESULT 8	AF276997	2699 bp	mRNA	linear	ROD 22-JUN-2000
LOCUS	Rattus norvegicus	tricarboxylate carrier-like protein mRNA,			
DEFINITION	complete cds.				
ACCESSION	AF276997				
VERSION	AF276997.1	GI:8650525			
KEYWORDS					
SOURCE	Rattus norvegicus (Norway rat)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;				
	Rattus.				
REFERENCE	1 (bases 1 to 2699)				
AUTHORS	Mashima,H., and Kojima,I.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-JUN-2000) Cell Biology, Institute for Molecular and Cellular Regulation, Gunma University, 3-39-15, Shova-machi, Maebashi, Gunma 371-8512, Japan				
FEATURES	Location/Qualifiers				
Source	1..2699				
	/organism="Rattus norvegicus"				
	/mol_type="mRNA"				
	/strain="Sprague Dawley"				
	/db_xref="taxon:10116"				
	233..1198				
	/note="CTP-like protein"				
	/codon_start=1				
	/product="tricarboxylate carrier-like protein"				
	/protein_id="AAF8249.1"				
	/db_xref="GI:8650526"				
	/translation="MGDLPLPINIOEPRMDSPFLGARHFFTYTDPRNLISGKOLEASRIINYVAGVTPRLTLDLARLYVDSDARHPTEGKYVLIGMSAQVPNNMTITGCMILFYRKTPPYVMQYNOSFNALVNTSNRGADPTVVODLGVTSATIGAVALTALGSGLTKHLPLVGRFVPPAAVAANNCINIPLMKRELDVGIVPVDEGQRGHSHVTAAGXGIFGVVYSRIGMAIPAMAIPVINMTRLEKKDKLRPMWLGALQGLVGFCLVFATPICALFEPPORSSIHVTRELPETRAIOIAOKPSIDVVVVYVNMKL"				
	2661..2666				
	polya_signal				
BASE COUNT	611 a	747 c	684 g	657 t	
GININ					
Query Match	53.9%	Score 664.2;	DB 10;	Length 2699;	
Best Local Similarity	82.5%;	Pred. No. 1.2e-139;			
Matches 785:	Conservative 0;	Mismatches 163;	Indels 3;	Gaps 2;	
QY	282	GTGGCTGGATTTCACCAATCCCCTGTGCCAACCGGGGTGGGGCGGCGGGAAGACTCTGCC	341		
Df	56	GTCCCGGAACCTCTGCCAAGAGGGGGGTGGGTGACAGCCAACCAAGATTTAGGAGACGCCCGCC	115		
QY	342	CTTCCCTGCTGTGTCGGCGCTCACGGCTGACGTCGCCGCTGTATGCTGGAGGGCCCGGCGG	401		

CDS

83..1048
/gene="sfm3"
/note="sfm3, similar to the flexed-tail (f) protein
sideroflexin 1 (sfm3)"
/codon_start=1
/product="sideroflexin 3"
/protein_id="AAK39430.1"
/db_xref="GI:13785616"
/translation="MGDPLNINIQEPRMDSITLGRARHFTVTPDRIILLSEGLE
ASRNIVONRAGVATPGLTEDOLMRARYVDSAPHTGEKVLIGRMSAPVPMNMTI
TGMLETFRKPTVFMQWNSFNALVNSRSGDAITVQQLGTAVSATGAVAT
ALGKSLPLKPLVGRFVPAVAANNCINIPMRBEOVGAPVDEAGORHSHV
TAAGKGLFOVYISRIGALPAMAIPIVIMNTLEKDELKRPMTGLQVGLVFCFLV
FAPPLCCALPORSISIVTRLEPRLRIQANQNSIVVYINKEK"
BASE COUNT 591 a 755 c 647 g 711 t 6 others
ORIGIN

Query Match 53.2%; Score 655.6; DB 10; Length 2710;
Best Local Similarity 87.7%; Pred. No. 1e-137;
Matches 726; Conservative 3; Mismatches 92; Indels 7; Gaps 1;

QY 412 GCGAGAGAGAGAGCGGTTCTGAGAGCTT-----CAGAGAGGATGAGAGAGAGAGATG 464
DB 26 RGCAGTGAAGAGAGCGGCTGAGAGCTTCTGAGAGCTTGAAGAGAGAGAGAGAGAGATG 85
QY 465 GGTGAATTCCTTTAGACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 524
DB 86 GGTGAATTCCTTTAGACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 145
QY 525 GCGAGAGAGAGAGAGCGGTTCTGAGAGCTT-----CAGAGAGGATGAGAGAGAGAGATG 584
DB 146 GCGAGAGAGAGAGAGCGGTTCTGAGAGCTTCTGAGAGCTTGAAGAGAGAGAGAGAGATG 205
QY 585 CAGTGAAG 644
DB 206 CAGTGAAG 265
QY 645 ATCAACGAG 704
DB 266 CTCACTGAG 325
QY 705 ACAG 764
DB 326 ACAG 385
QY 765 ATCAACGAG 824
DB 386 ATCACTGAG 445
QY 825 GTGAATTCCTTTAGACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 884
DB 446 GTGAATTCCTTTAGACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 505
QY 885 ACTGTGAG 944
DB 506 ACTGTGAG 565
QY 945 CTGAG 1004
DB 566 CTGAG 625
QY 1005 GCAG 1064
DB 626 GCAG 685
QY 1065 GTGAG 1124
DB 686 GTGAG 745
QY 1125 AAGCAG 1184
DB 746 AAGCAG 805

QY 1185 ATCCACACCATGATCATGACATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1232
DB 806 ATCCACACCATGATCATGACATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 853

RESULT 12
LOCUS BC005743
DEFINITION Mus musculus sideroflexin 1, mRNA (CDNA clone MGC:11926
IMAGE:3599636), complete cds.
ACCESSION BC005743
VERSION BC005743.1 GI:13543137
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Strausberg, R.L., Felingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Kaha, S.S., Loquiano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fehy, J., Holton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Boutard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Gilmwood, J., Schmutz, R.M.,
Butcherfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,
Scherer, A., Schein, J.E., Jones, S.J., and Merra, M.A.

TITLE
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932

REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgaps-remail.nih.gov

CDNA Library Preparation: Life Technologies, Inc.
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Louised, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nannavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 16 Row: P Column: 6.

FEATURES

source
1..1204
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:11926 IMAGE:3599636"

```

/tissue-type="Mammary tumor. Metallothionien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone-lib="NCI CGAP_Mam1"
/lab-host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 1204
/gene="Sfxn1"
/note="synonym: 2810002005R1k"
/db_xref="LocusID:14057"
/db_xref="MGI:2137677"
57..1025
CDS
/codon_start=1
/product="sideroflexin 1"
/protein_id="AAH05743.1"
/db_xref="GI:13543138"
/db_xref="GI:13785612"
/translation="MSGEVPPNINIKREPRMDSFGRASHFTVDPRLITNEOL
ENARKVVDYRQGIYPAGLTENELMRKAYTDSAPHPTGKMLIGMSAOVPMNT
ITGCMFTFRTTPAVLFQWQINOSFNNAVNTNSGAPLVNDELGTAVSATGAVA
TALGLNLTKRVSPLIGFVFPFAVAANNCINIPLMRRELKVGIPYDEGTRIGES
TNAKQAIQVVISRIILMAAPGMAIPPIINLTLEKALFKRPMMASAPIQTVLVGFL
VFATPLCCALFPKSSMSVTSLEDELQASIQRTHEIRRYVENKGL"
BASE COUNT      307 a      309 c      322 g      266 t
ORIGIN
Query Match      35.1%; Score 432; DB 10; Length 1204;
Best Local Similarity 72.7%; Pred. No. 4.1e-87;
Matches 558; Conservative 0; Mismatches 210; Indels 0; Gaps 0;
QY 465 GGTGAATTCCTTTAGACATCAATCCAGAGACCTGCTGGACCAAGTACTTCTCTG 524
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 63 GGGGAAGTCCACCCACCATTAACATCAAGAGCGCTCATGTGGACAGACAGCTTCA 122
QY 525 GGCAGAGCCCGGCACTTTTCACTGTGTACTGATCCTCGAAATCGCTGTCGCGGGA 584
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 123 GCGCGAGCGATCACTTCTTACGCGTTACTGATCCAGAAATATCTTTTAACGACGA 182
QY 585 CAGCTGAGGCTTCGGAACATCGTGAGAGTACAGAGCGCGCGGTGTGACCCGCA 644
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 183 CAGCTAGAGATGCGAGAGAAAGTGTCATGACTACAGCAAGAAATGTTCCGCGCG 242
QY 645 ATCAACCGAGACAGCTGTGAGAGCGCAAGTATGTATGACTCCGCTTCCATCCG 704
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 243 CTCACGGAATAATGTATGAGAGCGAAGTACCGTACGACTGCGCTTCCATCCAT 302
QY 705 ACAGGGGAGAGTGTGCTGATGTGGCCGATGTACGCGCAGGTGCGCCATGAAATG 764
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 303 ACCGCTGAAGAATGACTCTGATAGGAAGATGTACGTCAGTCCGATGAACATGACC 362
QY 765 ATCACTGGCTGATGCTTCAATCTTACAGAAAGACCCCAACCGTGTCTGTGCA 824
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 363 ATCAAGAGCTGATGATGATGCTTACCGGACCAACAGCGGTGTGCTTCTGGCA 422
QY 825 GTGAATGATGCTTCAATGCTATGTATGATCTTCAACACCGCATGTGACATCC 884
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 423 ATTAACCAAGTCTTCAATGCTGTGCTCACTACCAACCAACACCGGTGACGCT 482
QY 885 ACTGTAGAGAGCTGGGAGACAGCTATGTGATGAGCCACATGAGAGCTGTGGCC 944
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 483 ACTGTAGATGAGCTGGGAGAGCGCTTATGTTTCTGACACATGCGCGCTGTGCA 542
QY 945 CTGGAGACTCAATCCCTACCAAGACCTGCGCCCTTGTGCGAGATTTTGCCCT 1004
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 543 CTAGAGCTCAATGCTTAAACCAAGCGTGTCTCCACAGATGAGACGTTTGTCC 602
QY 1005 GCACAGTGGAGAGTGGCAATGATCAACATCCGCCCTGATGAGAGCAGAGAGCTG 1064
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 603 GCTCTGATGCTGTGCTAACTGATTAACCTGATGATGAGACAAAGGAGCTGAG 662
QY 1065 GTGGGATCCCGGTGCTGATGAGAGAGCTGAGAGCTGCTGCTGCTGCTGCTG 1124
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 663 GTTGTATTCCTGTCATGATGAAATGGACCTGCTTGGAGAAATGACCAACGACGA 722

```

```

QY 1125 AAGCAGGGAATCTTCAGGTGGTATTTCAAGAAATCTGCATGCGCATCTCCATG 1184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 723 AAGCAGGCAATCAGCAGAGTGTCTATCTTCAGAGATCTTCATGAGCCCGCGG 782
QY 1185 ATCCACGACATGATGATGACATCTGGAGAAAGAACTTCTCTAAG 1232
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 783 ATCCCTCGTTATCATGAAACACCTTGGAAGAAAGCCCTTTTAAAG 830
RESULT 13
AF325260      2803 bp      mRNA      linear      ROD 25-APR-2001
LOCUS
DEFINITION
Mus musculus sideroflexin 1 (Sfxn1) mRNA, complete cds; nuclear
gene for mitochondrial product.
ACCESSION
AF325260
AF325260.1 GI:13785611
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
MUS MUSCULUS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus.
REFERENCE
Fleming,M.D., Campagna,D.R., Haslett,J.N., Tremor,C.C. III and
Andrews,N.C.
A mutation in a mitochondrial transmembrane protein is responsible
for the pleiotropic hematological and skeletal phenotype of
flexed-tail (f/f) mice
Genes Dev. 15 (6), 652-657 (2001)
21172735
JOURNAL
MEDLINE
PUBMED
11274051
REFERENCE
Fleming,M.D., Campagna,D.R. and Andrews,N.C.
Direct Submission
Submitted (01-DEC-2000) Pathology, Children's Hospital, 300
Longwood Avenue, Boston, MA 02115, USA
FEATURES
source
1..2803
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/chromosome="13"
1..2803
/gene="Sfxn1"
78..1046
/gene="Sfxn1"
/note="Sfxn1"
/note="Sfxn1: mitochondrial membrane protein"
/codon_start=1
/product="sideroflexin 1"
/protein_id="AAK39428.1"
/db_xref="GI:13785612"
/translation="MSGEVPPNINIKREPRMDSFGRASHFTVDPRLITNEOL
ENARKVVDYRQGIYPAGLTENELMRKAYTDSAPHPTGKMLIGMSAOVPMNT
ITGCMFTFRTTPAVLFQWQINOSFNNAVNTNSGAPLVNDELGTAVSATGAVA
TALGLNLTKRVSPLIGFVFPFAVAANNCINIPLMRRELKVGIPYDEGTRIGES
TNAKQAIQVVISRIILMAAPGMAIPPIINLTLEKALFKRPMMASAPIQTVLVGFL
VFATPLCCALFPKSSMSVTSLEDELQASIQRTHEIRRYVENKGL"
122..123
variation
122..123
/gene="Sfxn1"
/note="Insertion mutation results in the flexed-tail (f)
mouse"
/replace="aat"
BASE COUNT      784 a      644 c      646 g      729 t
ORIGIN
Query Match      35.1%; Score 432; DB 10; Length 2803;
Best Local Similarity 72.7%; Pred. No. 3.9e-87;
Matches 558; Conservative 0; Mismatches 210; Indels 0; Gaps 0;
QY 465 GGTGAATTCCTTTAGACATCAATCCAGAGACCTGCTGGACCAAGTACTTCTCTG 524
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 84 GGGGAAGTCCACCCACCATTAACATCAAGAGCGCTCATGTGGACAGACAGCTTCA 143
QY 525 GGCAGAGCCCGGCACTTTTCACTGTGTACTGATCCTCGAAATCTGCTGCTGCGGGA 584

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2003, 08:15:20 ; Search time 365 Seconds
(without alignments)
9111.528 Million cell updates/sec

Title: US-09-990-415A-1

Perfect score: 1232
Sequence: 1 cccttaggcgcgcagggacag.....agaagaagacttcctgaag 1232

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

N_Geneseq_19Jun03.*
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1232	100.0	1232	22	ABN83754	Human protein clus
2	1231	99.9	1826	22	AAE27703	Human transport pr
3	1208.2	98.1	3047	22	AAH27272	Human cervical can
4	944	76.6	1729	22	AAI59123	Human polynucleoti
5	880	71.4	1061	24	ABN83755	Human protein clus
6	869.4	70.6	1375	24	ABO98865	Human ORF672 codin
7	848.8	68.9	2702	22	AAK94713	Human full-length
8	818.6	66.4	984	24	ABL39805	Human NS cDNA sequ

9	805.8	65.4	843	21	AAE75798	Human ORF1353
10	659	53.5	752	22	AAK91763	Human cDNA 5'-end
11	659	53.5	752	22	AAK93199	Human cDNA clone 1
12	632.2	51.3	2215	23	AAK84362	DNA encoding novel
13	516.6	41.9	7263	22	AAI60909	Human polynucleoti
14	399.4	32.4	2229	24	ABN83757	Human protein clus
15	398.2	32.3	1076	22	AAI59569	Human polynucleoti
16	398.2	32.3	1485	22	AAI26555	Human breast cance
17	398.2	32.3	2053	22	AAI61354	Human polynucleoti
18	398.2	32.3	2367	22	AAI21995	Human transporters
19	398.2	32.3	2468	21	AAE75794	Human ORF1349
20	396.8	32.2	1134	24	ABN84365	Alzheimer's diseas
21	396.6	32.2	2947	22	AAH44832	Rat cDNA encoding
22	385.4	31.3	1503	22	ABK54156	cDNA encoding huma
23	351.2	28.5	2766	22	AAH15610	Human cDNA sequenc
24	346.4	28.1	1084	24	ABK96715	DNA encoding human
25	335.4	27.2	1567	24	ABN83756	Human protein clus
26	335.4	27.0	1772	24	AAI33666	Human TRICH-21 cDN
27	332.2	27.0	2312	22	AAI59647	Human polynucleoti
28	332.2	27.0	2376	22	AAI57861	Human polynucleoti
29	330.6	26.8	2458	24	ABK35352	Human cDNA encodin
30	328.6	26.5	413	23	AAK84361	DNA encoding novel
31	315	25.6	984	24	ABK96716	DNA encoding spllic
32	315	25.6	984	24	ABN84366	Alzheimer's diseas
33	302.8	24.6	2642	22	AAH17727	Human cDNA sequenc
34	302.4	24.5	1575	24	ABL99804	Human secretory po
35	288.4	23.4	741	22	AAH06089	Human cDNA clone (
36	278.6	22.6	1286	23	ABL18135	Drosophila melanog
37	276.8	22.5	1548	24	ABL90501	Human polynucleoti
38	243.8	19.8	507	23	ABV45979	Human prostate exp
39	232.2	18.8	421	23	AAV16181	Human prostate exp
40	225.6	18.3	524	22	AAH31167	Human diagnostic a
41	222.8	18.1	806	22	AAH04784	Human cDNA clone (
42	173.6	14.1	4438	23	ABL11786	Drosophila melanog
43	170.6	13.8	1585	23	ABL11787	Drosophila melanog
44	164.2	13.3	1244	20	AAK97798	Extended human sec
45	162	13.1	2673	24	AAI36304	Human transporter

ALIGNMENTS

RESULT 1	ABN83754	standard; cDNA: 1232 BP.
ID	ABN83754	
AC	ABN83754	
XX	02-SEP-2002 (first entry)	
XX	Human protein cluster I gene.	
XX	Protein cluster I; human; metabolic disorder; obesity; diabetes;	
KW	antidiabetic; diagnosis; therapy; chromosome 10; gene; ss.	
XX	Homo sapiens.	
OS		
XX		
FX		
FT	Key	Location/Qualifiers
FT	CDS	450..1232
FT		/*tag= a
FT	/product=	"Protein cluster I"
FT	/partial	
FT	/note=	"the CDS does not include a stop codon"
PN	WO200242374-A1.	
XX		
PD	30-MAY-2002.	
XX		
XX	22-NOV-2001; 2001WO-SE02581.	
PF		
XX		
XX	24-NOV-2000; 2000SE-0004325.	
PR		
XX		
PA	(PHAA) PHARMACIA AB.	

XX Attersand A:
 PI
 XX
 DR WPI: 2002-500277/53.
 DR P-PSDB: ABB76445.
 XX
 PT Novel nucleic acid molecule encoding Protein Cluster I, useful in the
 PT diagnosis of metabolic diseases, such as obesity and diabetes, and in
 PT the identification of agents useful in the treatment of the diseases
 XX
 PS Claim 1: Page 23-24; 34pp; English.
 XX
 CC The present invention relates to the identification of a human gene
 CC family (see ABN83754-57) encoding a group of polypeptides (see
 CC ABB76445-48) referred to as Protein Cluster I. This family of
 CC homologous proteins was identified by an 'all-versus-all' BLAST
 CC procedure using all *Caenorhabditis elegans* proteins in the
 CC Wormpep20 database release. Proteins of unknown function were
 CC compared to the *Drosophila melanogaster* flybase database, and
 CC non-annotated protein clusters, conserved in both *C. elegans* and *D.*
 CC *melanogaster*, were used in a BLAST procedure against the Celera
 CC Human Genome Database, and Protein Cluster I proteins of unknown
 CC function were selected for study. The human part of Protein
 CC Cluster I comprises polypeptides encoded by 3 genes; an alternative
 CC splicing of the present sequence (deletion of nucleotides 624-794)
 CC gives the sequence in ABN83755. The present gene was identified in
 CC a human DNA sequence from clone Rpl1-108L7 on chromosome 10. It
 CC is expressed primarily in the nervous system and digestive system.
 CC The claimed Protein Cluster I nucleic acid molecules and proteins
 CC are proposed to be useful for differential identification of the
 CC tissue(s) or cell type(s) present in a biological sample, for
 CC diagnosis of diseases and disorders, including metabolic disorders
 CC and immune disorders, especially obesity and diabetes, and for
 CC identifying agents useful in the treatment of such diseases. The
 CC nucleic acid molecules are also useful as hybridisation probes, for
 CC chromosome and gene mapping, in PCR technologies, and in the
 CC production of sense or antisense nucleic acids.
 CC
 XX
 SO Sequence 1232 BP; 239 A; 368 C; 375 G; 250 T; 0 other:
 Query Match 100.0%; Score 1232; DB 24; Length 1232;
 Best Local Similarity 100.0%; Pred. No. 1.3e-303;
 Matches 1232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 421 GAAGGCGGTTCTGAGACTTCAGAGACGATGAGAAACAAATGGGTAATGGCTTTAG 480
 QY 481 ACATCAATCAACAGAACTCGCTGGAGCCAAAGTACTTTCTGGGAGAGCCGGCACT 540
 DB 481 ACATCAATCAACAGAACTCGCTGGAGCCAAAGTACTTTCTGGGAGAGCCGGCACT 540
 QY 541 TTTTCACTGTTACTGATCTCGAAATCTGCTGTCTGGGGGACACCTGGAACTTCTC 600
 DB 541 TTTTCACTGTTACTGATCTCGAAATCTGCTGTCTGGGGGACACCTGGAACTTCTC 600
 QY 601 GGAAATCGTGCAGAACTACAGGCGCGGTGGTCCAGGATCCAGAGACCAAGC 660
 DB 601 GGAAATCGTGCAGAACTACAGGCGCGGTGGTCCAGGATCCAGAGACCAAGC 660
 QY 661 TGTGGAGGCGCAAGTATGTATGATCTCCGCTTCCATCCGAGACAGGGAGAGTTGG 720
 DB 661 TGTGGAGGCGCAAGTATGTATGATCTCCGCTTCCATCCGAGACAGGGAGAGTTGG 720
 QY 721 TCTGATTTGGCCGATCTCAGCCAGGTGCCATGAACATGACATGCTGTCATGC 780
 DB 721 TCTGATTTGGCCGATCTCAGCCAGGTGCCATGAACATGACATGCTGTCATGC 780
 QY 781 TCACATTTCTACAGGAAGACCCCAACGCGTGGTCTGGCAGTGGGTGAATGTCATTCA 840
 DB 781 TCACATTTCTACAGGAAGACCCCAACGCGTGGTCTGGCAGTGGGTGAATGTCATTCA 840
 QY 841 ATGCCATTTGTAATCTACTCCAAACCGCAGTGTGACACTCCATCTGTGAGCAGCTGG 900
 DB 841 ATGCCATTTGTAATCTACTCCAAACCGCAGTGTGACACTCCATCTGTGAGCAGCTGG 900
 QY 901 GGACAGCCTTATGTAGTGCACACCACTGAGAGCTGTGGCCAGCGCCCTGGACTCAAATCC 960
 DB 901 GGACAGCCTTATGTAGTGCACACCACTGAGAGCTGTGGCCAGCGCCCTGGACTCAAATCC 960
 QY 961 TCACCAAGCACTGCCCCCTTGGTGGGCAAGATTGTGCTTGGACAGAGTGGCAGCTG 1020
 DB 961 TCACCAAGCACTGCCCCCTTGGTGGGCAAGATTGTGCTTGGACAGAGTGGCAGCTG 1020
 QY 1021 CCAACTGCATCAACATCCCTCGATGAGGAGAGAGACTGACAGTGGGATCCCGGTGG 1080
 DB 1021 CCAACTGCATCAACATCCCTCGATGAGGAGAGAGACTGACAGTGGGATCCCGGTGG 1080
 QY 1081 CTGATGAGGAGAGTCAAGATCTGCATGGCGATTCCTGCGATGAGGAGGAACTTCTCC 1140
 DB 1081 CTGATGAGGAGAGTCAAGATCTGCATGGCGATTCCTGCGATGAGGAGGAACTTCTCC 1140
 QY 1201 TGGACACTCTGGAGAGAAAGACTTCTGTAAG 1232
 DB 1201 TGGACACTCTGGAGAGAAAGACTTCTGTAAG 1232

RESULT 2
 AAF27703
 ID AAF27703 standard; cDNA; 1826 BP.
 XX
 AC AAF27703;
 DF 28-MAR-2001 (first entry)
 XX
 DE Human transport protein TPPT-3 coding sequence.
 XX
 KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
 KW neurological disorder; cardiovascular disorder; reproductive disorder;
 KW immune disorder; cancer; SS.
 XX
 OS Homo sapiens.
 XX
 PN W0200078953-A2.

XX 28-DEC-2000.
 PD 16-JUN-2000; 2000WO-US16668.
 XX
 XX
 XX 17-JUN-1999; 99US-0139923.
 PR 10-AUG-1999; 99US-0148177.
 PR 18-AUG-1999; 99US-0149357.
 PR 28-OCT-1999; 99US-0162287.
 XX
 PA (INCYTE GENOMICS INC.
 PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
 PI Baughn MR, Azimzal Y, Lu DM, Au-Young J, Patterson C;
 XX WPI: 2001-041424/05.
 DR P-PSDB; AAB60083.
 XX
 XX Isolated polypeptide with a human transport protein sequence is useful
 PT for the diagnosis, prevention and treatment of disorders associated
 PT with the immune, reproductive and cardiovascular systems -
 XX
 XX Claim 5; Page 142-143; 165pp; English.
 XX
 CC The present invention provides the protein and coding sequences for 43
 CC novel human transport proteins (designated TPTs). These can be used in
 CC the diagnosis and treatment of transport, metabolic, neurological,
 CC reproductive, cardiovascular and immune disorders, and cell proliferative
 CC disorders such as cancer.
 XX
 XX Sequence 1826 BP; 415 A; 519 C; 512 G; 380 T; 0 other;
 XX
 Query Match 99.9%; Score 1231; DB 22; Length 1826;
 Best Local Similarity 100.0%; Pred. No. 2,6e-303;
 Matches 1231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 541 TTTCACGTGTACTGATCTCTGAAATCTGCTGCTGCCGGGACAGCTGGAAAGCTTCTCG 600
 |||
 Qy 602 GAACATCGTGCAGAACTACAGGGCCGGGTGGTGAACCCAGGAGATCACCAGGACACT 661
 |||
 Db 601 GAACATCGTGCAGAACTACAGGGCCGGGTGGTGAACCCAGGAGATCACCAGGACACT 660
 |||
 Qy 662 GTGAGGGCCAAAGTATGTATGACTCCGCTTCATCCGAGACAGAGGGGAAGTGGT 721
 |||
 Db 661 GTGAGGGCCAAAGTATGTATGACTCCGCTTCATCCGAGACAGAGGGGAAGTGGT 720
 |||
 Qy 722 CTTGATTTGGCCGATGTAGCCAGGTGCCATGAAATGACATCACTGCTGATGCT 781
 |||
 Db 721 CTTGATTTGGCCGATGTAGCCAGGTGCCATGAAATGACATCACTGCTGATGCT 780
 |||
 Qy 782 CACATTTACAGGAAGACCCCAACCGTGGTCTTGGAGGGGGAATCAAGTCTTAA 841
 |||
 Db 781 CACATTTACAGGAAGACCCCAACCGTGGTCTTGGAGGGGGAATCAAGTCTTAA 840
 |||
 Qy 842 TGCCATTTGTTAACTACTCCAAACCGCAGTGTGACACTCCATCACTGTAGAGCACTGGG 901
 |||
 Db 841 TGCCATTTGTTAACTACTCCAAACCGCAGTGTGACACTCCATCACTGTAGAGCACTGGG 900
 |||
 Qy 902 GACAGCTATGTAGTGTCCACACTGAGCTGTGGCCACGGCCCTGGACTCAAAATCCT 961
 |||
 Db 901 GACAGCTATGTAGTGTCCACACTGAGCTGTGGCCACGGCCCTGGACTCAAAATCCT 960
 |||
 Qy 962 CACCAAGACCTGCCCCCTGTGGCGGAGATTTGGCCCTTTCAGAGAGTGGCACTGC 1021
 |||
 Db 961 CACCAAGACCTGCCCCCTGTGGCGGAGATTTGGCCCTTTCAGAGAGTGGCACTGC 1020
 |||
 Qy 1022 CAACTGATCAACATCCCTGTATGAGGACAGAGAGCTGAGGTGGCATCCCGGTGGC 1081
 |||
 Db 1021 CAACTGATCAACATCCCTGTATGAGGACAGAGAGCTGAGGTGGCATCCCGGTGGC 1080
 |||
 Qy 1082 TGATGAGGACAGGTAGAGGCTTGGCTACTGGTACTGTGACACCAAGAGGAATCTTCA 1141
 |||
 Db 1081 TGATGAGGACAGGTAGAGGCTTGGCTACTGGTACTGTGACACCAAGAGGAATCTTCA 1140
 |||
 Qy 1142 GGTGATTTCAAGAAATCTGCATGAGGATTCCTGCCATGCCATCCACCATGATCAT 1201
 |||
 Db 1141 GGTGATTTCAAGAAATCTGCATGAGGATTCCTGCCATGCCATCCACCATGATCAT 1200
 |||
 Qy 1202 GGACACTCTGGAGAAAGAACTTCTGAG 1232
 |||
 Db 1201 GGACACTCTGGAGAAAGAACTTCTGAG 1231
 |||

RESULT 3
 AAH72722/C
 ID AAH72722 standard; CDNA: 3047 BP.
 XX
 XX AAH72722;
 AC
 XX
 DE 19-SEP-2001 (first entry)
 XX
 DE Human cervical cancer marker nucleic acid 3996.
 XX
 DE Cervical cancer: cytostatic; pre-malignant condition; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200142467-A2.
 PD 14-JUN-2001.
 XX
 XX 08-DEC-2000; 2000WO-US33312.
 PR 08-DEC-1999; 99US-0169681.
 PR 21-DEC-1999; 99US-0171350.
 PR 14-MAR-2000; 2000US-0189315.
 PR 12-MAY-2000; 2000US-0203791.
 PR 09-JUN-2000; 2000US-0210600.

PR 21-JUL-2000; 2000US-0220114.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA Schlegel R, Deeds J, Berger A, Zhao X;
 PI WPI: 2001-375006/39.
 XX
 XX New isolated nucleic acid for diagnosing and treating cervical cancer
 PT and for assessing and detecting compounds for treating the cancer -
 XX
 PS Claim 1: Page 815-816; 1051ip; English.
 XX
 CC The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful: to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;
 CC and to select and/or assess the efficacy of a compound or therapy for
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy.
 XX
 XX Sequence 3047 BP; 629 A; 790 C; 891 G; 733 T; 4 other;
 SO
 Query Match 98.1%; Score 1208.2; DB 22; Length 3047;
 Best Local Similarity 99.8%; Pred. No. 2e-297;
 Matches 1210; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 20 GCCGAGCGTACCTGTCGCCGCGGAGCGAGTCTTACCCAGCCAGTTCTGCTGCA 79
 DB 3035 GTCCGCTTACCTGTCGCCGCGGAGCGAGTCTTACCCAGCCAGTTCTGCTGCA 2976
 QY 80 CGCCCTAGCTATTCGCAATTTAGGCTTGGTCTGCTGCTTCCCTCCGCTCGAA 139
 DB 2975 CGCCCTAGCTATTCGCAATTTAGGCTTGGTCTGCTGCTTCCCTCCGCTCGAA 2916
 QY 140 CCACCTCTCTCTGAGCGGAGCCAGCTACCGGGGCTCTGGAATTCGACACCCCTCGG 199
 DB 2915 CCACCTCTCTCTGAGCGGAGCCAGCTACCGGGGCTCTGGAATTCGACACCCCTCGG 2856
 QY 200 GCACCTTTCGAGGCTCCGTCGAGGAGCTCAAGGGGAGAGCGGAGCTGAGCTGATT 259
 DB 2855 GCACCTTTCGAGGCTCCGTCGAGGAGCTCAAGGGGAGAGCGGAGCTGAGCTGATT 2796
 QY 260 TGGTCAGGGGCTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 319
 DB 2795 TGGTCAGGGGCTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2736
 QY 320 GCGGCGCGGAGGAGCTCTGCGCCCTCCCTGCTGTCGTCGTCGTCGTCGTCGTCGTCG 379
 DB 2735 GCGGCGCGGAGGAGCTCTGCGCCCTCCCTGCTGTCGTCGTCGTCGTCGTCGTCGTCG 2676
 QY 380 GATGGCTGGAGAGCCCGCGCGGAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 439
 DB 2675 GATGGCTGGAGAGCCCGCGCGGAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2616
 QY 440 TCAGAGAGCGANTGGAAGCAAAATGGTGAATTCCTTGAATTCGATCAATCCAGGAACC 499
 DB 2615 TCAGAGAGCGANTGGAAGCAAAATGGTGAATTCCTTGAATTCGATCAATCCAGGAACC 2556
 QY 500 TCGCTGGGACCAAGTACTTCTCTGGGAGAGCCCGGACACTTTTCACTGTTACTGATCC 559
 DB 2555 TCGCTGGGACCAAGTACTTCTCTGGGAGAGCCCGGACACTTTTCACTGTTACTGATCC 2496
 QY 560 TCGAAATCTGCTGTCGCGGAGACAGCTGGAAGCTTCTGGAACATCTGTCGGAACATA 619
 DB 2495 TCGAAATCTGCTGTCGCGGAGACAGCTGGAAGCTTCTGGAACATCTGTCGGAACATA 2436
 QY 620 CAGGCGCGGAGTGTGACCCAGGAGATCACGAGAGACAGTGTGGAGGCGCAAGTATGT 679
 DB 2435 CAGGCGCGGAGTGTGACCCAGGAGATCACGAGAGACAGTGTGGAGGCGCAAGTATGT 2376
 QY 680 GATGACTCCGCTTCATCCGACACAGGAGGAGAGTGTGATTCGATTCGATTCGATTCG 739

DB 2375 GATGACTCCGCTTCATCCGAGACAGGGGAGAGTGTGCTGATTCGATTCGATTCG 2316
 QY 740 AGCCAGGTGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 799
 DB 2315 AGCCAGGTGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2256
 QY 800 CCACACCGTGTGTTCTGTCGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATG 859
 DB 2255 CCACACCGTGTGTTCTGTCGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATG 2196
 QY 860 CAACCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 919
 DB 2195 CAACCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2136
 QY 920 CACCATGAGCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 979
 DB 2135 CACCATGAGCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2076
 QY 980 CTTGGTGGGAGATTTGTCCTTTGACAGAGTGGAGCTGCAATGCAATGCAATGCAATGCAAT 1039
 DB 2075 CTTGGTGGGAGATTTGTCCTTTGACAGAGTGGAGCTGCAATGCAATGCAATGCAATGCAAT 2016
 QY 1040 CCTGATGAGGAGAGAGAGTGTGAGTGGGATGATGATGATGATGATGATGATGATGATGATG 1099
 DB 2015 CCTGATGAGGAGAGAGAGTGTGAGTGGGATGATGATGATGATGATGATGATGATGATGATG 1956
 QY 1100 GCTTGGCTACTGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1159
 DB 1955 GCTTGGCTACTGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1896
 QY 1160 CTGATGAGGAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1219
 DB 1895 CTGATGAGGAGATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1836
 QY 1220 AGACTTCCTGAG 1232
 DB 1835 AGACTTCCTGAG 1823
 RESULT 4
 ID AA159123
 AA159123 standard; cDNA; 1729 BP.
 AC AA159123;
 XX
 XX 22-OCT-2001 (first entry)
 DE
 XX Human polynucleotide seq ID NO 1326.
 KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 PN
 XX WO200153312-A1.
 PD 26-JUL-2001.
 XX
 XX 26-DEC-2000; 2000WO-US34263.
 XX
 XX 21-JAN-2000; 2000US-0488725.
 XX 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 DR WPI: 2001-442253/47.
 DR P-PSDB: AAM39967.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS Claim 1; SEQ ID NO 1326; 10078bp; English.
 XX The invention relates to human nucleic acids (A157798-A161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with neoptotic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC XX
 SQ Sequence 1729 BP; 406 A; 469 C; 491 G; 363 T; 0 other;
 Query Match 76.6%; Score 944; DB 22; Length 1729;
 Best Local Similarity 100.0%; Pred. No. 3e-230;
 Matches 944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 829 ATGATCCTTCAATGCGCATTTGTTACTACTCAACCGCAGTGTGACACTCCATCACTG 888
 DB 626 ATGATGCTCTCAATGCGCATTTGTTACTACTCAACCGCAGTGTGACACTCCATCACTG 885
 QY 889 TGAGGACGTGGGGACAGCCTATGTAGTGTGCCACTGAGCTGTGGCCAGGCGCTGG 948
 DB 686 TGAAGGACGTGGGGACAGCCTATGTAGTGTGCCACTGAGCTGTGGCCAGGCGCTGG 745
 QY 949 GACTCAATCCCTCACCAGCAGCTGCGCCCTTGGTGGGAGATTGTGCTTGTGACG 1008
 DB 746 GACTCAATCCCTCACCAGCAGCTGCGCCCTTGGTGGGAGATTGTGCTTGTGACG 805
 QY 1009 CAGTGGCAGCTGGCCAACTGATCAATCCCTGATGAGGACAGAGAGCTGCGAGTGG 1068
 DB 806 CAGTGGCAGCTGGCCAACTGATCAATCCCTGATGAGGACAGAGAGCTGCGAGTGG 865
 QY 1069 GCATCCCGGTGCTGATGAGGACAGTGAAGGCTTGCTACTGCTGATGCTGACCAAGC 1128
 DB 866 GCATCCCGGTGCTGATGAGGACAGTGAAGGCTTGCTACTGCTGATGCTGACCAAGC 925
 QY 1129 AGGGAATCTTCCAGGTGGTGTATTCAGAAATCTGCATGGGAGTCTGCCATGGCCATTC 1188
 DB 926 AGGGAATCTTCCAGGTGGTGTATTCAGAAATCTGCATGGGAGTCTGCCATGGCCATTC 985
 QY 1189 CACCACTGATCATGAGACACTGTGAGAAAGAAAGCTTCTGTAAG 1232
 DB 986 CACCACTGATCATGAGACACTGTGAGAAAGAAAGCTTCTGTAAG 1029

RESULT 5
 ABN83755
 ID ABN83755 standard; cDNA; 1061 BP.
 XX
 AC ABN83755;
 XX
 DT 02-SEP-2002 (first entry)
 XX
 DE Human protein cluster I gene.
 XX
 KW Protein cluster I; human; metabolic disorder; obesity; diabetes;
 KW antidiabetic; diagnosis; therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 450..680
 FT /*tag- a
 FT /product- "Protein cluster I"
 FT /partial
 FT /note- "the CDS does not include a stop codon"
 PN W0200242324-A1.
 XX
 PD 30-MAY-2002.
 XX
 PF 22-NOV-2001; 2001WO-SE02581.
 XX
 PR 24-NOV-2000; 2000SE-0004325.
 XX
 PA (PHAA) PHARMACIA AB.
 XX
 PI Attersand A;
 DR WPI: 2002-500277/53.
 DR P-PSDB: ABB76446.
 XX
 PT Novel nucleic acid molecule encoding Protein Cluster I, useful in the
 PT diagnosis of metabolic diseases, such as obesity and diabetes, and in
 PT the identification of agents useful in the treatment of the diseases
 XX
 PS Claim 1; Page 25-26; 34pp; English.

CC The present invention relates to the identification of a human gene
 CC family (see ABN83754-57) encoding a group of polypeptides (see
 CC ABN76445-48) referred to as Protein Cluster I. This family of
 CC homologous proteins was identified by an 'all-versus-all' BLAST
 CC procedure using all *Caenorhabditis elegans* proteins in the
 CC Wormpep20 database release. Proteins of unknown function were
 CC compared to the *Drosophila melanogaster* Flybase database, and
 CC non-annotated protein clusters, conserved in both *C. elegans* and *D.*
 CC *melanogaster*, were used in a BLAST procedure against the Celera
 CC Human Genome database, and Protein Cluster I proteins of unknown
 CC function were selected for study. The human part of Protein
 CC Cluster I comprises polypeptides encoded by 3 genes. The present
 CC sequence is an alternatively spliced version of the sequence given
 CC in ABN83754. It is expressed primarily in the male genitalia.
 CC The claimed Protein Cluster I nucleic acid molecules and proteins
 CC are proposed to be useful for differential identification of the
 CC tissue(s) or cell type(s) present in a biological sample, for
 CC diagnosis of diseases and disorders, including metabolic disorders
 CC and immune disorders, especially obesity and diabetes, and for
 CC identifying agents useful in the treatment of such diseases. The
 CC nucleic acid molecules are also useful as hybridisation probes, for
 CC chromosome and gene mapping, in PCR technologies, and in the
 CC production of sense or antisense nucleic acids.

XX Sequence 1061 BP; 203 A; 318 C; 322 G; 218 T; 0 other;

Query Match 71.4%; Score 880; DB 24; Length 1061;

Best Local Similarity 86.1%; Pred. No. 5e-214; Mismatches 0; Indels 171; Gaps 1;

Matches 1061; Conservativity 0; Mismatches 0; Indels 171; Gaps 1;

QY 1 CCCTTGGGCGCCAGGACACCCGCTTACCTGCTCCGGGACGGAGATTCTTTACCC 60
 DB 1 CCTTGGGCGCCAGGACACCCGCTTACCTGCTCCGGGACGGAGATTCTTTACCC 60
 QY 61 ACCCAATTGTTGTTGAGCCCTAGCTATTCGGAATTTAGGCTTGGCTGCT 120
 DB 61 ACCCAATTGTTGTTGAGCCCTAGCTATTCGGAATTTAGGCTTGGCTGCT 120
 QY 121 TGTTCCTCCCTCGGCTGGAACCACTCTTCTGTGACCGGACGACTACCGGGCTCTGG 180
 DB 121 TGTTCCTCCCTCGGCTGGAACCACTCTTCTGTGACCGGACGACTACCGGGCTCTGG 180
 QY 121 TGTTCCTCCCTCGGCTGGAACCACTCTTCTGTGACCGGACGACTACCGGGCTCTGG 180
 DB 121 TGTTCCTCCCTCGGCTGGAACCACTCTTCTGTGACCGGACGACTACCGGGCTCTGG 180
 QY 181 AATTGCCACCCCTCCCTGCGGACCCCTTGAAGGCTTCCCTGCTGCTGCTGCTGCTGCT 240
 DB 181 AATTGCCACCCCTCCCTGCGGACCCCTTGAAGGCTTCCCTGCTGCTGCTGCTGCTGCT 240
 QY 241 CGGAGCTGAGCTGAGTTGCTGCAAGCGCTGCTGTGTGAGGCTTGTGCTGCTGCTGCT 300
 DB 241 CGGAGCTGAGCTGAGTTGCTGCAAGCGCTGCTGTGTGAGGCTTGTGCTGCTGCTGCT 300
 QY 301 CCCCCTGCGCCACCGGGGCTGGGCGGCGGAGAGCTCTGCTGCTGCTGCTGCTGCTGCT 360
 DB 301 CCCCCTGCGCCACCGGGGCTGGGCGGCGGAGAGCTCTGCTGCTGCTGCTGCTGCTGCT 360
 QY 361 CACGCGTGAAGTCCCGGCTGATGCTGAGGAGCGGCGGCGGAGAGAGAGAGAGAGAG 420
 DB 361 CACGCGTGAAGTCCCGGCTGATGCTGAGGAGCGGCGGCGGAGAGAGAGAGAGAGAG 420
 QY 421 GAAGGCGGTTCTGAGAGCTTCAAGAGCGATGAGAAAGCAAAATGGGTGAATTCCTTTAG 480
 DB 421 GAAGGCGGTTCTGAGAGCTTCAAGAGCGATGAGAAAGCAAAATGGGTGAATTCCTTTAG 480
 QY 481 ACATCAACATCCAGGAACCTGCTGGAGCAAAAGTACTTCTTCTGGGAGAGCCCGGCACT 540
 DB 481 ACATCAACATCCAGGAACCTGCTGGAGCAAAAGTACTTCTTCTGGGAGAGCCCGGCACT 540
 QY 541 TTTTCACTGTTACTGATCTCGAAATCTGCTGCTCGGAGGACAGACTGGAAGCTTCTC 600
 DB 541 TTTTCACTGTTACTGATCTCGAAATCTGCTGCTCGGAGGACAGACTGGAAGCTTCTC 600
 QY 601 GGAACATCTGCAAGAACTACAGGCGCGCTGTGTGACCCAGGAGTACACGAGAGACAGC 660
 DB 601 GGAACATCTGCAAGAA----- 616

QY 661 TGTGAGGCGCAAGATATGTATGACTCCGCTTCATCCGAGACAGAGGAGAGGTGG 720
 DB 617 ----- 616
 QY 721 TCTGATTTGGCCCATGTCAGCCAGGTGCCATGAACATGACCATTACTGCTGATGC 780
 DB 617 ----- 616
 QY 781 TCACATTTCTACGAGAGACCCCAACGCTGCTTGTGCGATGGGTGAATCAGTCTTCA 840
 DB 617 -----CTACAGAGAGACCCCAACGCTGCTTGTGCGATGGGTGAATCAGTCTTCA 669
 QY 841 ATGCATTTGTTAACTACTCCCAACCGCAGTGTGACACTCCCATCTGTAAGGACGTGG 900
 DB 670 ATGCATTTGTTAACTACTCCCAACCGCAGTGTGACACTCCCATCTGTAAGGACGTGG 729
 QY 901 GGACAGCTTATGATGATGCCACCACTGAGACTGTGGCCAGAGGCTTGGACTCAATCCC 960
 DB 730 GGACAGCTTATGATGATGCCACCACTGAGACTGTGGCCAGAGGCTTGGACTCAATCCC 789
 QY 961 TCACCAAGCACTGCCCCCTTGTGCGAGATTGTGCTTGTGACAGTGGACAGCTG 1020
 DB 790 TCACCAAGCACTGCCCCCTTGTGCGAGATTGTGCTTGTGACAGTGGACAGCTG 849
 QY 1021 CCAACTGCATCAACATCCCTGATGAGGAGAGAGAGCTGCAAGTGGGATCCGGTGG 1080
 DB 850 CCAACTGCATCAACATCCCTGATGAGGAGAGAGAGCTGCAAGTGGGATCCGGTGG 909
 QY 1081 CTGATGAGGAGCTGAGAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 DB 910 CTGATGAGGAGCTGAGAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 969
 QY 1141 AGGTGATGATTAAGATCTGATGAGGAGATTCCTGCTGATGAGGATCCACACTGATCA 1200
 DB 970 AGGTGATGATTAAGATCTGATGAGGAGATTCCTGCTGATGAGGATCCACACTGATCA 1029
 QY 1201 TGGACACTCTGAGAGAGAAAGACTTCTGTAAG 1232
 DB 1030 TGGACACTCTGAGAGAGAAAGACTTCTGTAAG 1061

RESULT 6
 ABQ98865
 ID ABQ98865 standard; DNA; 1375 BP.
 AC ABQ98865;
 XX
 DT 04-NOV-2002 (first entry)
 XX
 DE Human ORF672 coding sequence.
 XX
 KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vascular;
 KW Antiinflammatory; gene therapy; human; ORF; atherosclerotic; platelet;
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
 KW cancer; cardiovascular disease; allergy; autoimmune disease;
 KW wound healing; blood coagulation disorder; inflammatory disorder; ds.
 XX
 OS Homo sapiens.
 XX
 PN US2002082206-A1.
 XX
 PD 27-JUN-2002.
 XX
 PF 30-MAY-2001; 2001US-0867550.
 XX
 PR 30-MAY-2000; 2000US-208427P.
 XX
 PA (LEACH) LEACH M D.
 PA (MEHR) MEHRABAN F.
 PA (CONL) CONLEY P B.
 PA (TOPP) TOPPER J N.
 PA (LAMD) LAM D.


```

Db      61 ATGGGTAAATGTCCTTTAGACATACATCCAGAACCTCCCTGGGACCAAGTACTTTC
Qy      522 CTGGGACAGCCCGGACATTTTTCATGTTACTGATATCCGAAATCTGCTGCGGG
Db      121 CTGGGACAGAG-CGCGACATTTTTCATGTTACTGATATCCGAAATCTGCTGCGGG
Qy      582 GCACAGCTGGAAGCTTCTCGGACATCGTGAGAACTACAGAGGCGCGGCTGTGACCCCA
Db      180 GCACAGCTGGAAGCTTCTCGGACATCGTGAGAACTACAGAGGCGCGGCTGTGACCCCA
Qy      642 GGGATCACCCGAGGACAGCTGTGAGAGGCGCAAGTATGTTGACTCGGCTTCATCCG
Db      240 GGGATCACCCGAGGACAGCTGTGAGAGGCGCAAGTATGTTGACTCGGCTTCATCCG
Qy      702 GACACAGGGGAGAGGAGTGTCTCTGATTTGGCCGACATGTCAGCCAGTCCCATGAACATG
Db      300 GACACAGGGGAGAGGAGTGTCTCTGATTTGGCCGACATGTCAGCCAGTCCCATGAACATG
Qy      762 ACATACACTGCTGCTCATGCTCATATTTACAGAGAACCCCAACCGTGTGCTGCGAG
Db      360 ACATACACTGCTGCTCATGCTCATATTTACAGAGAACCCCAACCGTGTGCTGCGAG
Qy      822 TGGGTGAATGATGCTTCAATGCAATGTTTAACTACTTCAACCCGAGTGTGACATCCC
Db      420 TGGGTGAATGATGCTTCAATGCAATGTTTAACTACTTCAACCCGAGTGTGACATCCC
Qy      882 ATCACTGTGAGGACAGCTGTGGGACAGCCATATGTGATGTCACACACTGGAGTGTGCGCAG
Db      480 ATCACTGTGAGGACAGCTGTGGGACAGCCATATGTGATGTCACACACTGGAGTGTGCGCAG
Qy      942 GCCCTGGGACTCAAAATCCCTCACCAGACACCTGCCCTTGGTGGCAGATTTGTGCC
Db      540 GCCCTGGGACTCAAAATCCCTCACCAGACACCTGCCCTTGGTGGCAGATTTGTGCC
Qy      1002 TTTCAGACAGTGGAGTGGCAACTGCAATCAACATCCCTCGATGAGGACAGAGAGCTG
Db      600 TTTCAGACAGTGGAGTGGCAACTGCAATCAACATCCCTCGATGAGGACAGAGAGCTG
Qy      1062 CAGGTGGGACATCCCGGTGGCTGATGATGAGGACAGAGGCTGTGCTACTCGTACTGCA
Db      660 CAGGTGGGACATCCCGGTGGCTGATGATGAGGACAGAGGCTGTGCTACTCGTACTGCA
Qy      1122 GCCAAGAGGAAATCTTCCAGAGTGTGATTTCAAGAAATCTGCATGGCGATTCGCCATG
Db      720 GCCAAGAGGAAATCTTCCAGAGTGTGATTTCAAGAAATCTGCATGGCGATTCGCCATG
Qy      1182 GCCATCCACACATGATCATGAGACATCTGGAGAAAGAACTCTCGAAG
Db      780 GCCATCCACACATGATCATGAGACATCTGGAGAAAGAACTCTCTGGAAG

```

RESULT 9
AAC75798
ID AAC75798 standard; cDNA: 843 BP.
AC AAC75798;

08-FEB-2001 (first entry)

Human ORFX ORF1353 polynucleotide sequence SEQ ID NO:2705.

Human: open reading frame; ORFX; detection; cytostatic; hepatotropic;
vulnereary; antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
anticonvulsant; osteoporotic; antidiarrheal; immunosuppressant; candidant;
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
hypotensive; dermatological; immunosuppressive; antineoplastic;
antiviral; antibacterial; antifungal; antineumatic; antithyroid;
antianemic; gene therapy; cancer; proliferative disorder; hypertension;
neurodegenerative disorder; osteoarthritis; graft vs host disease;
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
cholesterol ester storage; systemic lupus erythematosus; infection;
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;

allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
bone damage; cartilage damage; antiinflammatory disease; coagulation;
thrombosis; contraceptive; ss.
Homo sapiens.
MO200058473-A2.
05-OCT-2000.
31-MAR-2000; 2000WO-US08621.
31-MAR-1999; 99US-0127607.
02-APR-1999; 99US-0127636.
05-APR-1999; 99US-0127728.
30-MAR-2000; 2000US-0540763.
(CURA-) CURAGEN CORP.
Shinkets RA, Leach M;
WPI: 2000-602362/57.
P-PSDB: AAB41589.
Novel nucleic acids and peptides derived from open reading frame X,
useful for treating e.g. cancers, proliferative disorders,
neurodegenerative disorders and cardiovascular disease -
Claim 5; Page 1942; 5507pp; English.
AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
which represent the human ORFX open reading frames 1 to 3161. The ORFX
sequences have activities such as: cytostatic; hepatotropic; vulnereary;
antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
anticonvulsant; antidiarrheal; antineoplastic; immunosuppressant;
immunostimulant; cancer; thrombolytic; coagulant; vasotropic;
antidiabetic; hypotensive; dermatological; immunosuppressive;
antineumatic; antithyroid; antifungal; antineumatic; antithyroid;
antianemic; gene therapy; cancer; proliferative disorder; hypertension;
neurodegenerative disorder; osteoarthritis; graft vs host disease;
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; viral;
cholesterol ester storage; systemic lupus erythematosus; infection;
severe combined immunodeficiency (SCID); AIDS; viral;
cholesterol ester storage; systemic lupus erythematosus; infection;
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
coagulation; to inhibit thrombosis; and as a contraceptive.
Sequence 843 BP; 183 A; 241 C; 252 G; 165 T; 2 other:
Query Match 65.4%; Score 805.8; DB 21; Length 843;
Best Local Similarity 97.4%; Pred. No. 3.5e-195;
Matches 819; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

Qy      362 ACGGCTGACGTCGCCGCGATGAGTGGGAGGCGCGGGGAGCAGCGAGAGAGG
Db      3 ACGGCTGACGTCGCCGCGATGAGTGGGAGGCGCGGGGAGCAGCGAGAGAGG
Qy      422 AAGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTCCTTTAGA
Db      63 AAGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTGAATTCCTTTAGA
Qy      482 CATCAATATCCAGAGAACCTGCTGTGGACCAAAAGTACTTCTCTGGGACAGCCCGGCACTT
Db      123 CATCAATATCCAGAGAACCTGCTGTGGACCAAAAGTACTTCTCTGGGACAGCCCGGCACTT
Qy      542 TTTCAGTGTACTGATCTCTGAAATCTGCTGTGGGACAGCTGGAAGGCTTCTCG
Db      183 TTTCAGTGTACTGATCTCTGAAATCTGCTGTGGGACAGCTGGAAGGCTTCTCG

```



```

Db      661 CCCTGATGAGCAAAANAACTGCAGAGTGGCATCCCAATGGCTGATGAGGACAGTC 720
QY      1096 AGAGGCTTGCTACTCGCT 1114
Db      721 AAAAGCTTGCTACTCGCT 739

RESULT 11
AAK93199 standard; cDNA; 752 BP.
XX
AC      AAK93199;
XX
DT      06-NOV-2001 (first entry)
XX
DE      Human cDNA clone representative sequence, SEQ ID NO: 1659.
XX
KW      Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS      Homo sapiens.
XX
PN      EP1130094-A2.
XX
PD      05-SEP-2001.
XX
PF      07-JUL-2000; 2000EP-0114089.
XX
PR      08-JUL-1999; 99JP-0194486.
XX
PR      11-JAN-2000; 2000JP-0118774.
XX
PR      02-MAY-2000; 2000JP-0183765.
XX
PA      (HELI-) HELIX RES INST.
XX
PI      Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX
PI      Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR      WPI; 2001-524255/58.
XX
PT      830 Primers useful for synthesizing full length cDNA clones and their
XX
PT      use in genetic manipulation -
XX
PS      Example 11; SEQ ID NO 1659; 1380bp + sequence listing; English.
XX
CC      The invention relates to primers for synthesizing full length cDNA
CC      clones. 830 cDNA molecules encoding a human protein have been
CC      isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC      molecules have been determined. Primers for synthesizing the full length
CC      cDNA are useful for clarifying the function of the protein encoded by
CC      the cDNA. The full length clones were obtained by construction of full
CC      length enriched cDNA libraries that were synthesised by the oligo-capping
CC      method. The primers enable the production of the full length cDNA easily
CC      without any special methods. The present sequence was used as the
CC      representative sequence from a human clone which was used in
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ      Sequence 752 BP; 173 A; 208 C; 214 G; 149 T; 8 other:

Query Match      53.5%; Score 659; DB 22; Length 752;
Best Local Similarity 96.5%; Pred. No. 8e-158;
Matches 713; Conservative 0; Mismatches 21; Indels 5; Gaps 4;

QY      381 ATGGCTGGAGGGCCCGGGCGACAGCGGAGCGAGAGAGAGCGGTTCTGAGAGCTT 440
Db      1 ATGGCTGGAGGGCCCGGGCGACAGCGGAGCGAGAGAGAGAGCGGTTCTGAGAGCTT 60
QY      441 CAGAGAGCGATGGAAGCAAAATGGGTGAATTCCTTTAGACATCAACATCCAGAACT 500
Db      61 CAGAGAGCGATGGAAGCAAAATGGGTGAATTCCTTTAGACATCAACATCCAGAACT 120
QY      501 CGCTGGGACCAAGTACTTCTCTGGGAGAGCCCGGACACTTTTCACTGTACTGATCCT 560

```

```

Db      121 CGCTGGGACCAAGTACTTCTCTGGGAGAGCCCGGACACTTTTCACTGTACTGATCCT 180
QY      561 CGAATGTGCTGCTCCGGGGGACACCTGGAACTCTGGGAACATCGGCAACTAC 620
Db      181 CGAATGTGCTGCTCCGGGGGACACCTGGAACTCTGGGAACATCGGCAACTAC 240
QY      621 AGGGCCGGCTGTGAGCCCA-GGGATCACCAGAGACACCTGTGAGGGCCAAATGT 679
Db      241 AGGGCCGGCTGTGAGCCCAAGGAGATCACCAGAGACACCTGTGAGGGCCAAATGT 300
QY      680 GTATGACTCCGCTTCCATCCGAGACAGGGAGAGGTGCTCTGATTGGCCGATGTC 739
Db      301 GTATGACTCCGCTTCCATCCGAGACAGGGAGAGAGGTGCTCTGATTGGCCGATGTC 360
QY      740 AGCCGAGTCCCATGAAATGACATGACATGACATGCTGCTGATGCTCATTTACAGAGAG 799
Db      361 AGCCGAGTCCCATGAAATGACATGACATGACATGCTGCTGATGCTCATTTACAGAGAG 420
QY      800 CCCAACCCTGTTCTGGCAGTGGGTAATCACTCCCTTCATGCTTAACTACTC 859
Db      421 CCCAACCCTGTTCTGGCAGTGGGTAATCACTCCCTTCATGCTTAACTACTC 480
QY      860 CAACCCGAGTGTGACACTCCATCACTGTGAGGACAGCTGGGACACCTATGTAGTGC 919
Db      481 CAACCCGAGTGTGACACTCCATCACTGTGAGGACAGCTGGGACACCTATGTAGTGC 540
QY      920 CACACCTGAGCTGTGGCCAGAGCCCTGGGAGTCAATCCCTCCAGAGACCTGGCCCC 979
Db      541 CACACCTGAGCTGTGGCCAGAGCCCTGGGAGTCAATCCCTCCAGAGACCTGGCCCC 600
QY      980 CTTGGTCGAGATTTGTGCTTGGCAGAGTGGCAGTGGCCAAAC-TGCATCAACATCC 1038
Db      601 CTTGGTCGAGATTTGTGCTTGGCAGAGTGGCAGTGGCCAAAC-TGCATCAACATCC 660
QY      1039 CCCTGATGAGGACAGAGAGACTGC-AGTGGCATCCCGGTGC--CTGATGAGGACAGTC 1095
Db      661 CCCTGATGAGGACAAANAACTGCAGAGTGGCATCCCAATGGCTGATGAGGACAGTC 720
QY      1096 AGAGGCTTGCTACTCGCT 1114
Db      721 AAAAGCTTGCTACTCGCT 739

RESULT 12
AAS84362
AAS84362 standard; cDNA; 2215 BP.
XX
AC      AAS84362;
XX
DT      13-FEB-2002 (first entry)
XX
DE      DNA encoding novel human diagnostic protein #20166.
XX
DE      DNA encoding novel human diagnostic protein #20166.
XX
KW      Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW      food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS      Homo sapiens.
XX
PN      W0200175067-A2.
XX
PD      11-OCT-2001.
XX
PF      30-MAR-2001; 2001WO-US08631.
XX
PR      31-MAR-2000; 2000US-0540217.
XX
PR      23-AUG-2000; 2000US-0649167.
XX
PA      (HYSE-) HYSEO INC.
XX
PI      Drmanac RT, Liu C, Tang YF;
XX
DR      WPI; 2001-639362/73.

```



```

Db      468 AAGTGGAGACGACCCCTCACTGTCATAGAGTGGGAACAGCTTACGTTCTGCAACAC 527
QY      926 TGAAGCTGTGGCCACGGCCCTGGGAGCTCAAAATCCCTCACCAGACCTGCCCCCTTGGT 985
Db      528 TGGTGGCGGTAGCAACAGCTGTAGAGCTCAATGACATGACCAAGCATGTCTCACCACGTAT 587
QY      986 CGGCAGATTGTGGCCCTTTTGCAGCAGTGGCAGCTGCCAAGTGCATCAACATCCCCCTGAT 1045
Db      588 AGGACGTTTGTGCTTCCCTTGTGCTGAGCTGCTGAATGCAATTAATATTCATTAAT 647
QY      1046 GAGCGAGAGAGAGCTGAGAGTGGGAGTCCCGGTGGCTGATGAGGACAGTGCAGAGCTTGG 1105
Db      648 GAGGCANAAGGAACCTCAAGTTGGCATTTCCGTCACGAGTGAATGGGAACCGCTTGGG 707
QY      1106 CTACTCGGTGACTGACAGCAAGCAGGAGATCTTCAGGTGGTGAATTCAGAAATCTGCAT 1165
Db      708 GGAAGTCGGCGAAGCTGCGAAGCAAGGCATACGCAAGTGTCTGTCAGGATTCAT 767
QY      1166 GCGGATTCTGCGCATGCGCATCCACCACTGATCATGAGACACTCTGAGAGAGAAAGACTT 1225
Db      768 GGCAGCCCCCTGGCATGGCCATCCCTCCATTCATATGACACACTTGGAAAAAGAAAGCCTT 827
QY      1226 CTTGAAG 1232
Db      828 TTGAAG 834

```

Search completed: October 16, 2003, 14:00:18
 Job time : 372 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2003, 13:54:01 ; Search time 360 Seconds
(without alignments)
8987.194 Million cell updates/sec

Title: US-09-990-415A-1

Perfect score: 1232
Sequence: 1 cccttaggcgcgcagagcag.....agaagaagacttcgtgaag 1232

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1232	100.0	1232	10	US-09-990-415A-1
2	944	76.6	1729	14	US-10-037-270-1016
3	880	71.4	1061	10	US-09-990-415A-3
4	869.4	70.6	1375	9	US-09-867-550-1343
5	399.4	32.4	2269	10	US-09-990-415A-7
6	398.2	32.3	2978	14	US-10-198-846-9767
7	396.8	32.2	1134	14	US-10-014-338-1
8	335.4	27.2	1567	10	US-09-990-415A-5
9	332.2	27.0	2376	13	US-10-098-841-64
10	330.6	26.8	2458	9	US-09-822-849A-490
11	315	25.6	984	14	US-10-014-338-3
12	272.8	22.1	670	14	US-10-206-901B-41
13	212.4	17.2	489	11	US-09-918-995-23576
14	197.2	16.0	293	10	US-09-783-590-2110
15	166.2	13.5	3342	12	US-09-814-353-20153
16	131.2	10.6	522	11	US-09-918-995-5675

17	122	9.9	648	14	US-10-198-846-7595	Sequence 7595, Ap
18	118.2	9.6	496	14	US-10-106-698-2663	Sequence 2663, Ap
19	89	7.2	533	11	US-09-918-995-26900	Sequence 26900, A
20	73.4	6.0	238	11	US-09-991-936-1857	Sequence 1857, Ap
21	46.4	3.8	248	10	US-09-728-445-404	Sequence 404, Ap
22	41.6	3.4	2303	10	US-09-880-107-3849	Sequence 3849, Ap
23	41.6	3.4	2303	11	US-09-922-146-3	Sequence 146, Ap
24	40.6	3.3	135638	12	US-10-314-657-1	Sequence 1, App11
25	40	3.2	2631	14	US-10-156-761-6208	Sequence 6208, Ap
26	40	3.2	9025608	14	US-10-156-761-1	Sequence 1, App1
27	39.8	3.2	5025	10	US-09-960-253-176	Sequence 176, App
28	39.8	3.2	42999	10	US-09-799-462A-17	Sequence 17, App1
29	39.8	3.2	42999	11	US-09-836-911A-17	Sequence 17, App1
30	39.8	3.2	42999	12	US-09-738-630-73	Sequence 73, App1
31	39.8	3.2	42999	13	US-10-125-767-17	Sequence 17, App1
32	39.8	3.2	42999	14	US-10-151-081-17	Sequence 17, App1
33	39.8	3.2	42999	14	US-10-287-313-17	Sequence 17, App1
34	39.8	3.2	42999	14	US-10-219-694-17	Sequence 17, App1
35	39.4	3.2	390	14	US-10-156-761-44	Sequence 44, App1
36	39.4	3.2	9025608	14	US-10-156-761-1	Sequence 1, App1
37	38.8	3.1	2561	10	US-09-976-740-48	Sequence 48, App1
38	38.8	3.1	2561	13	US-10-023-529-48	Sequence 48, App1
39	38.8	3.1	2561	13	US-10-023-523-48	Sequence 48, App1
40	38.4	3.1	6453	12	US-09-960-706-618	Sequence 618, App
41	38	3.1	2000	10	US-09-887-576-874	Sequence 874, App
42	37.6	3.1	1874	10	US-09-323-998D-33	Sequence 33, App1
43	37.4	3.0	838	13	US-10-027-632-163967	Sequence 163967, App
44	37.4	3.0	838	13	US-10-027-632-163968	Sequence 163968, App
45	37.4	3.0	1683	14	US-10-156-761-5724	Sequence 5724, Ap

ALIGNMENTS

RESULT 1
US-09-990-415A-1
Sequence 1, Application US/09990415A
Patent No. US20020165182A1
GENERAL INFORMATION:
APPLICANT: Pharmacia AB
TITLE OF INVENTION: Protein Cluster I
FILE REFERENCE: 00349
CURRENT APPLICATION NUMBER: US/09/990,415A
CURRENT FILING DATE: 2001-11-21
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 1232
TYPE: DNA
ORGANISM: human
FEATURE:
NAME/KEY: CDS
LOCATION: (450)..(1232)
US-09-990-415A-1

Query Match Best Local Similarity 100.0%; Score 1232; DB 10; Length 1232;
Matches 1232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CCCTTAGGCGCCAGGAGCGAGCGTACTGTCGGGCGAGCGAGTTCTTACC	60
Db	1	CCCTTAGGCGCCAGGAGCGAGCGTACTGTCGGGCGAGCGAGTTCTTACC	60
Qy	61	ACCCAGTTCTGTTCTGACGCCCTAGCTCAATTCGCAAAATTAGGCTTGCTGCT	120
Db	61	ACCCAGTTCTGTTCTGACGCCCTAGCTCAATTCGCAAAATTAGGCTTGCTGCT	120
Qy	121	TGTTCCCTCGGCTCGACACCACTCTTCTGAGCGAGCGAGCTTACGGGCTCTGG	180
Db	121	TGTTCCCTCGGCTCGACACCACTCTTCTGAGCGAGCGAGCTTACGGGCTCTGG	180
Qy	181	AATGACACCCCTCGGCGAGCGCTTGAAGCTCCGCGAGGAGCGTACAGGGGACAG	240
Db	181	AATGACACCCCTCGGCGAGCGCTTGAAGCTCCGCGAGGAGCGTACAGGGGACAG	240

```

Db      181 AATTGCCACCCCTCCTGGGACACCTTGAGGCTCCGCGAGAGGACGTCACGGGGCAGAG 240
QY      241 CGGGACCTGAGCCCTGATTTGCTGCAGAGCGCTCTGTGTGTGGCTGGCTTTCGCCAAT 300
Db      241 CGGGACCTGAGCCCTGATTTGCTGCAGAGCGCTCTGTGTGTGGCTGGCTTTCGCCAAT 300
QY      301 CCCGCTGCCACCGGGGTGGGGCGGGGAGAGTCCGTCGCCCTCCCGCTGGGTGGGGGT 360
Db      301 CCCGCTGCCACCGGGGTGGGGCGGGGAGAGTCCGTCGCCCTCCCGCTGGGTGGGGGT 360
QY      361 CACGCTGACGCTCCCGCTGTGTGTGGAGGGCCCGGGCGAGACGAGGAGAGAG 420
Db      361 CACGCTGACGCTCCCGCTGTGTGTGGAGGGCCCGGGCGAGACGAGGAGAGAG 420
QY      421 GAAGCGGCTCTGTGAGCTTCAGAGACGATGAGAAACAAATGGGTGAATGGCTTTAG 480
Db      421 GAAGCGGCTCTGTGAGCTTCAGAGACGATGAGAAACAAATGGGTGAATGGCTTTAG 480
QY      481 ACATCAACATCCAGAGACCTCGCTGGGACCAAACTACTTTCCTGGGAGAGCCGGGACT 540
Db      481 ACATCAACATCCAGAGACCTCGCTGGGACCAAACTACTTTCCTGGGAGAGCCGGGACT 540
QY      541 TTTTCACTGTACTGATCTCGAAATCTGCTGTCTCCGGGGCAGAGCTGGAGCTTCTC 600
Db      541 TTTTCACTGTACTGATCTCGAAATCTGCTGTCTCCGGGGCAGAGCTGGAGCTTCTC 600
QY      601 GGAACATCTGTGAGAACTACAGGGCGCGCTGTGTGACCCAGGAGTACCGAGACAGC 660
Db      601 GGAACATCTGTGAGAACTACAGGGCGCGCTGTGTGACCCAGGAGTACCGAGACAGC 660
QY      661 TGTGAGGGGCAAGTATGTATGATCTCCGGCTTCATCCGGGAGACAGGGGAGAGGTGG 720
Db      661 TGTGAGGGGCAAGTATGTATGATCTCCGGCTTCATCCGGGAGACAGGGGAGAGGTGG 720
QY      721 TCTGATTTGGCCGATGTACAGCCAGGTGCCCATGAACATGACCATCTGGCTGATGC 780
Db      721 TCTGATTTGGCCGATGTACAGCCAGGTGCCCATGAACATGACCATCTGGCTGATGC 780
QY      781 TCACATTTCTACAGGAAGAACCCCAACCGTGTCTTGGCAAGTGGGTGAATCAGTCTTCA 840
Db      781 TCACATTTCTACAGGAAGAACCCCAACCGTGTCTTGGCAAGTGGGTGAATCAGTCTTCA 840
QY      841 ATGCCATTTGTAATCTACCAACCGAGTGTGACACTCCCATCTGTGAGGAGAGCTGG 900
Db      841 ATGCCATTTGTAATCTACCAACCGAGTGTGACACTCCCATCTGTGAGGAGAGCTGG 900
QY      901 GGACAGGCTATGTGAGTGCACACACTGAGCTGTGGCCAGGCGCTGGACTCAAAATCCC 960
Db      901 GGACAGGCTATGTGAGTGCACACACTGAGCTGTGGCCAGGCGCTGGACTCAAAATCCC 960
QY      961 TCACCAAGCACTGCCCCCTTGGTGGGAGATTTGGCCCTTGGACAGAGTGGCACTG 1020
Db      961 TCACCAAGCACTGCCCCCTTGGTGGGAGATTTGGCCCTTGGACAGAGTGGCACTG 1020
QY      1021 CCAACTGCATCAACATCCCTCGATGAGGAGAGAGAGTGCAGTGGGATCCCGGTGG 1080
Db      1021 CCAACTGCATCAACATCCCTCGATGAGGAGAGAGAGTGCAGTGGGATCCCGGTGG 1080
QY      1081 CTGATGAGGAGAGTGCAGAGCTTGGCTACTCGGTGATGCAAGCCAGAGGAACTCTTCC 1140
Db      1081 CTGATGAGGAGAGTGCAGAGCTTGGCTACTCGGTGATGCAAGCCAGAGGAACTCTTCC 1140
QY      1141 AGGTGTGATTTCAAGAAATCGATGGGATTCCTGTCATGGCCATCCACCACTGATCA 1200
Db      1141 AGGTGTGATTTCAAGAAATCGATGGGATTCCTGTCATGGCCATCCACCACTGATCA 1200
QY      1201 TGGACACTGTGAGAGAAAGAACTTCTGAAG 1232
Db      1201 TGGACACTGTGAGAGAAAGAACTTCTGAAG 1232

```

RESULT 2
US-10-037-270-1016

```

: Sequence 1016, Application US/10037270
: Publication No. US20030104529A1
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhang, Jie
: APPLICANT: Ren, Feiyan
: APPLICANT: Chen, Rui-hong
: APPLICANT: Zhao, Qing A.
: APPLICANT: Weinman, Tom
: APPLICANT: Xue, Aldong J.
: APPLICANT: Yang, Yonghong
: APPLICANT: Wang, Jian-Rui
: APPLICANT: Zhou, Ping
: APPLICANT: Ma, Yungling
: APPLICANT: Wang, Dunrui
: APPLICANT: Wang, Zhilwei
: APPLICANT: Tillinghast, John
: APPLICANT: Drmanac, Radoje T.
: TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
: FILE REFERENCE: 784CIP2B
: CURRENT FILING DATE: 2002-01-04
: PRIOR FILING DATE: 2000-04-25
: PRIOR APPLICATION NUMBER: 09/488,725
: NUMBER OF SEQ ID NOS: 1104
: SOFTWARE: PL-Fl-Genes Version 1.0
: SEQ ID NO 1016
: LENGTH: 1729
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (247)..(1371)
: US-10-037-270-1016

Query Match      76.6%; Score 944; DB 14; Length 1729;
Best Local Similarity 100.0%; Pred. No. 1.2e-275;
Matches 944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      289 GCTTCTCCCAATCCCGCTGCCACCGGGTGGGGCCCGGGGAAGCTCTCCCTCCCT 348
Db      86 GGTTCGCCAATCCCGCTGCCACCGGGTGGGGCCCGGGGAAGCTCTCCCTCCCT 145
QY      349 GCTGTGGGGTCAAGGCTGACGTCGCCGCTGATGGCTGGAGAGGGCCGGCGGACAGC 408
Db      146 GCTGTGGGGTCAAGGCTGACGTCGCCGCTGATGGCTGGAGAGGGCCGGCGGACAGC 205
QY      409 GGAAGCAGAGAGGAAGCGGCTTCTGAGAGCTTCAGAGAGCGATGAAAGCAAAATGGGTG 468
Db      206 GGAAGCAGAGAGGAAGCGGCTTCTGAGAGCTTCAGAGAGCGATGAAAGCAAAATGGGTG 265
QY      469 AATTGCCCTTAGAATCAATCAATCAAGAACTCCCTGGAGCAAAAGTACTTCTTGGGCA 528
Db      266 AATTGCCCTTAGAATCAATCAATCAAGAACTCCCTGGAGCAAAAGTACTTCTTGGGCA 325
QY      529 GAGCCCGGCACTTTTCACTGTACTGATCCTCGAAATCTGCTCTCGGGGACAGC 588
Db      326 GAGCCCGGCACTTTTCACTGTACTGATCCTCGAAATCTGCTCTCGGGGACAGC 385
QY      589 TGGAACTTCTCGGAACATCGTGCAGAACTACAGAGCGCGGGGTGTACCCACAGATCA 648
Db      386 TGGAACTTCTCGGAACATCGTGCAGAACTACAGAGCGCGGGGTGTGTACCCACAGATCA 445
QY      649 CCGAGACAGCTGTGAGAGGCGCAAGTATGTATGATCGCTCGCTTCATCCGAGACAG 708
Db      446 CCGAGACAGCTGTGAGAGGCGCAAGTATGTATGATCGCTTCATCCGAGACAG 505
QY      709 GGAAGAGTGTGCTGATTTGGCCGATGTACAGCCAGTGGCCATGACATGACATCA 768

```


Db 506 GGGAGAGAGGTGCTCGATTGGCGCATGTCAGCCAGGTGCCATGAAATGACATCA 565
QY 769 CTGGTCATGCTCATCTTACAGAGAAAGCCCAACCGTGGTTCAGAGGGTGA 828
Db 566 CTGGTCATGCTCATCTTACAGAGAAAGCCCAACCGTGGTTCAGAGGGTGA 625
QY 829 ATCAGTCTTCAATGCGATTGTTAACTACTCAACCGCAGTGGTGACATCCCATCACTG 888
Db 626 ATCAGTCTTCAATGCGATTGTTAACTACTCAACCGCAGTGGTGACATCCCATCACTG 685
QY 889 TGAGGACAGCTGGGGAGACAGCTATGTAGTCCACACACTGGAGCTGTGGCCAGGCTTGG 948
Db 686 TGAGGACAGCTGGGGAGACAGCTATGTAGTCCACACACTGGAGCTGTGGCCAGGCTTGG 745
QY 949 GACTCAATCCCTCACCAAGACACTGCCCCCTTGGTGGCAGATTGTGCTCCCTTTCAG 1008
Db 746 GACTCAATCCCTCACCAAGACACTGCCCCCTTGGTGGCAGATTGTGCTCCCTTTCAG 805
QY 1009 CAGTGGCAGCTGCCAAGCTGCATCAATCCCTGATGAGGACAGAGAGCTGCAGGTGG 1068
Db 806 CAGTGGCAGCTGCCAAGCTGCATCAATCCCTGATGAGGACAGAGAGCTGCAGGTGG 865
QY 1069 GCATCCCGGTGGCTGATGAGGACAGGTGAGAGGCTGGCTACTGCTGATGACAGGCTAGC 1128
Db 866 GCATCCCGGTGGCTGATGAGGACAGGTGAGAGGCTGGCTACTGCTGATGACAGGCTAGC 925
QY 1129 AGGGAATCTTCCAGGTGGTATTTCAAGAAATCTGCATGGGATTCCTGCCATGCCATTC 1188
Db 926 AGGGAATCTTCCAGGTGGTATTTCAAGAAATCTGCATGGGATTCCTGCCATGCCATTC 985
QY 1189 CACACTGATCATGAGACACTCTGAGAGAAAGACTTCCTGAAG 1232
Db 986 CACACTGATCATGAGACACTCTGAGAGAAAGACTTCCTGAAG 1029

RESULT 3
US-09-990-415a-3
; Sequence 3, Application US/09990415A
; Patent No. US20020165182A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Protein Cluster I
; FILE REFERENCE: 00349
; CURRENT APPLICATION NUMBER: US/09/990,415A
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1061
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (450)..(680)
US-09-990-415a-3

Query Match 71.4%; Score 880; DB 10; Length 1061;
Best Local Similarity 86.1%; Pred. No. 2,6e-256;
Matches 1061; Conservative 0; Mismatches 0; Indels 171; Gaps 1;

QY 1 CCTTAGGCGCCAGGAGACAGGCTTACCTGGTCCCGGACCGGAGTTCTTTACCC 60
Db 1 CCTTAGGCGCCAGGAGACAGGCTTACCTGGTCCCGGACCGGAGTTCTTTACCC 60
QY 61 ACCCAAGTCTGGTTCGAGCGCCCTAGCTATCCGCAATTTAGGCTTGGGCTGGCT 120
Db 61 ACCCAAGTCTGGTTCGAGCGCCCTAGCTATCCGCAATTTAGGCTTGGGCTGGCT 120
QY 121 TGTTCCTCTCGGCTCGAAGACACTCTTCTGAGCGGACAGCTACCGGGGCTCCTGG 180
Db 121 TGTTCCTCTCGGCTCGAAGACACTCTTCTGAGCGGACAGCTACCGGGGCTCCTGG 180

QY 181 AATTGCCACCCCTCCCTGGGACACCTTGAGGCTTCCTGTGAGGAGCTACGCGGGCAGAG 240
Db 181 AATTGCCACCCCTCCCTGGGACACCTTGAGGCTTCCTGTGAGGAGCTACGCGGGCAGAG 240
QY 241 CGGAGCTGAGCCTGAGTTTGTCTGACAGGCGTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 300
Db 241 CGGAGCTGAGCCTGAGTTTGTCTGACAGGCGTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 300
QY 301 CCCGCTGCCACCGGGGGGCGCGCGGGGAAAGCTTCCTGCCCTCCCTGCTGGTGGCGCT 360
Db 301 CCCGCTGCCACCGGGGGGCGCGCGGGGAAAGCTTCCTGCCCTCCCTGCTGGTGGCGCT 360
QY 361 CACGCTGACGCTCCCGGTGATGCTGGAGGCGCGCGGGGACAGCGGAGAGAGAG 420
Db 361 CACGCTGACGCTCCCGGTGATGCTGGAGGCGCGCGGGGACAGCGGAGAGAGAGAG 420
QY 421 GAAGGCGTCTGAGACCTTCAAGAGACGATGGAAGCAAAATGGGTGAATTGGCTTAG 480
Db 421 GAAGGCGTCTGAGACCTTCAAGAGACGATGGAAGCAAAATGGGTGAATTGGCTTAG 480
QY 481 ACATCAACATCCAGGAACCTTCGCTGGAGCAAAAGTACTTTCTTGGGAGAGCCCGGACT 540
Db 481 ACATCAACATCCAGGAACCTTCGCTGGAGCAAAAGTACTTTCTTGGGAGAGCCCGGACT 540
QY 541 TTTTCACTGTACTGATCTCTGAAATCTGCTGCTGCTCGGGGACACAGCTGGAACCTTCTC 600
Db 541 TTTTCACTGTACTGATCTCTGAAATCTGCTGCTGCTCGGGGACACAGCTGGAACCTTCTC 600
QY 601 GGAACATCTGAGCAAACTACAGGCGCGGCGGTGGACCCCAAGGATCACGAGAGACAGC 660
Db 601 GGAACATCTGAGCAAA----- 616
QY 661 TGTGAGGCGCAAGTATGTATGATCTCGCCTTCATCCGAGACAGGGAGAGAGGTGG 720
Db 617 ----- 616
QY 721 TCCGTATGGCCCGCATCTCAGCCCAAGTGGCCATGAAATGACATCACTGGCTGCATGC 780
Db 617 ----- 616
QY 781 TCACATCTCAGGAAGCCCAACCGTGGTGTCTGGAGTGGGTGAATCAGTCTTCA 840
Db 617 -----CTACAGGAAGCCCAACCGTGGTGTCTGGAGTGGGTGAATCAGTCTTCA 669
QY 841 ATGCCATTTTAACTACTCTCAACCGCAGTGGTGAACCTCCATCACTGTGAGGAGCTGG 900
Db 670 ATGCCATTTTAACTACTCTCAACCGCAGTGGTGAACCTCCATCACTGTGAGGAGCTGG 729
QY 901 GGACAGCTATGTGAGTGGCCACCACTGGAGCTGTGGCCAGGCGCTGGGACTCAAAATCCC 960
Db 730 GGACAGCTATGTGAGTGGCCACCACTGGAGCTGTGGCCAGGCGCTGGGACTCAAAATCCC 789
QY 961 TCACAGAGCACTGCCCGCTTGGTGGGAGATTTGTGCCCTTGTGACAGTGGGACACTG 1020
Db 790 TCACAGAGCACTGCCCGCTTGGTGGGAGATTTGTGCCCTTGTGACAGTGGGACACTG 849
QY 1021 CCAACTGCAATCAACATCCCGCTGATGAGGAGAGAGACTGCAGAGTGGGCAATCCGGTGG 1080
Db 850 CCAACTGCAATCAACATCCCGCTGATGAGGAGAGAGAGACTGCAGAGTGGGCAATCCGGTGG 909
QY 1081 CTGATGAGGAGAGTCAAGAGCTTGGCTACTCGGTGATGTGACAGCAACAGAGGAATCTTCC 1140
Db 910 CTGATGAGGAGAGTCAAGAGCTTGGCTACTCGGTGATGTGACAGCAACAGAGGAATCTTCC 969
QY 1141 AGGTGATGATTTCAAGAAATCTGCATGCGATTCCTGCGATGGCCATCCACACTGATCA 1200
Db 970 AGGTGATGATTTCAAGAAATCTGCATGCGATTCCTGCGATGGCCATCCACACTGATCA 1029
QY 1201 TGGACACTGTGAGAGAAAGACTTCTGGAAG 1232
Db 1030 TGGACACTGTGAGAGAAAGACTTCTGGAAG 1061


```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1121)..(1122)
OTHER INFORMATION: where "n" is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1125)..(1126)
OTHER INFORMATION: where "n" is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1132)..(1133)
OTHER INFORMATION: where "n" is any nucleotide
US-10-014-338-1

```

```

Query Match      32.2%; Score 396.8; DB 14; Length 1134;
Best Local Similarity 69.8%; Pred. No. 7.3e-110;
Matches 536; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

QY 465 GGTGAATTCCTTTGACATCAACATCCAGAACCTGCTGGGAGCAAGTACTTCTCTG 524
DB 17 GGAGAACTACCAACCAATTTACATCAAGAACTCGATGGATCAAGCACTTTCAT 76
QY 525 GGCAGAGCCCGGCACTTTTCACTGTACTGTATCTCTGAATCTGCTGCTCGGGGCA 584
DB 77 GGACGAGCCCAATCATTTCTCACTGTACTGACCCAGAGACATTTCTGTACCAAGAA 136
QY 585 CAGCTGGAAGCTTCTCGAATCTGCGAATCAAGAGCCGCGCTGTGACCCAGGG 644
DB 137 CACCTGGAAGCTGCGAATTAATGATGATTAAGGCAAGAAATTTCTCTCTGT 196
QY 645 ATACCGAGAGCAGCTGTGAGGAGCAATGATGTATGATCTCCGCTTCATCCAGAC 704
DB 197 CTACGAAATGAAATGTGAGAGCAAAATGATGATGATGATGATGATGATGATGATG 256
QY 705 AAGAGGAGAGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 764
DB 257 ACTGGAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 316
QY 765 ATCACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 824
DB 317 ATCACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 376
QY 825 GTGAATCACTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 884
DB 377 ATTAACAGATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 436
QY 885 ACTGAGAGCAGCTGGGAGACCTGATGATGATGATGATGATGATGATGATGATGATG 944
DB 437 ACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 496
QY 945 CTGGAGCTCAATTCCTCAACAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1004
DB 497 CTAGAGCTCAATTCCTCAACAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
QY 1005 GCAGAGTGGAGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1064
DB 557 GCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 616
QY 1065 GTGGGATCCCGTGTGCTGATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1124
DB 617 GTGGGATCCCGTGTGCTGATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 676
QY 1125 AAGCAGAGGAATTCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1184
DB 677 AAGCAGAGGAATTCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 736
QY 1185 ATCCACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1232
DB 737 ATCCACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 784

```

```

Sequence 5, Application US/09990415A
Patent No. US20020165182A1
GENERAL INFORMATION:
APPLICANT: Pharmacia AB
TITLE OF INVENTION: Protein Cluster I
FILE REFERENCE: 00349
CURRENT APPLICATION NUMBER: US/09/990,415A
CURRENT FILING DATE: 2001-11-21
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 1567
TYPE: DNA
ORGANISM: human
FEATURE:
NAME/KEY: CDS
LOCATION: (47)..(1015)
US-09-990-415a-5

```

```

Query Match      27.2%; Score 335.4; DB 10; Length 1567;
Best Local Similarity 64.5%; Pred. No. 3.3e-91;
Matches 501; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 456 AGCAAAATGGGTAATGCTTTAGACATCAATCCAGAACTGCTGGGACCAAGT 515
DB 41 AGCAAGATGGAGCTGACCTGTCTGTGCTTTAATGATGATGATGATGATGATGATG 100
QY 516 ACTTCTGAGGAGAGCCCGGCACTTTTCACTGTGATGATGATGATGATGATGATG 575
DB 101 ACTTCTGAGGAGAGAGTGAAGCACTCTTAACATCAAGAGCCCGGCACTGTCTGTA 160
QY 576 TCCGGGAGCAGCTGGAACCTTCTGGAATCTGTCGAATCAATCAAGAGCCCGGCTG 635
DB 161 TCTGAGCGGAGCTGAGTGGCCAGGATGATGATGATGATGATGATGATGATGATG 220
QY 636 ACCCAGGAGATCAGGAGGACAGCTGTGAGGAGCCAAATGATGATGATGATGATG 695
DB 221 CCCAGGAGCAGGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATG 280
QY 696 CATCCGAGACAGGAGGAGAGTGTCTGATGATGATGATGATGATGATGATGATG 755
DB 281 CACCCGAGACAGGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATG 340
QY 756 AACATGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 815
DB 341 GGCATGATATATCAGGAGGCTTCAATGCTTCAATGATGATGATGATGATGATGATG 400
QY 816 TGGCAGTGGGTAATCAGTCTTCAATGCTTCAATGATGATGATGATGATGATGATG 875
DB 401 TGGCAGTGGGTAATCAGTCTTCAATGATGATGATGATGATGATGATGATGATGATG 460
QY 876 ATCCCATCATCTGTGAGGAGCTGAGGAGACCTATGATGATGATGATGATGATG 935
DB 461 TCCCCCATCATCTGTGAGGAGCTTCTTCAATGATGATGATGATGATGATGATGATG 520
QY 936 GCCAGGAGCTGGAGCTCAATTCCTCAACAGACACTGCTGCTGCTGCTGCTGCTGCT 995
DB 521 GCCAGGAGCTGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 580
QY 996 GTGCCCTTTGAGCAGATGAGCTGCAATCTCAATCAATCAATCAATCAATCAATCA 1055
DB 581 GTGCCCTTTGAGCAGATGAGCTGCAATCTCAATCAATCAATCAATCAATCAATCA 640
QY 1056 GAGCTGCAAGTGGGATCCCGTGTGATGAGGAGCTGATGATGATGATGATGATG 1115
DB 641 GAGCTGCAATGAGGATGCTGTAAGAGCAGAGAAATGAAATGAGATGATGATGATG 700
QY 1116 ACTGAGCCCAAGAGGAATTCAGAGTGTGATGATGATGATGATGATGATGATGATG 1175
DB 701 AAGAGCTGGGAGCAGTACCAAGATGATGATGATGATGATGATGATGATGATGATG 760
QY 1176 GCCATGGCCATCCACAGCTGATGATGATGATGATGATGATGATGATGATGATG 1232

```


	Matches	420; Conservative	0; Mismatches	175; Indels	0; Gaps	0;
QY	465	GGTGAATTCCTTTAGACATCAATCCAGAACCTGCTGGACCAAGTACTTCTG	524			
Db	17	GGAAACTACCAACCAACTTAACTACAGAACCTTCATGGGATCAAAAGCATTTCTT	76			
QY	525	GGCAGAGCCCGGCACTTTTTCAGCTTTCATCTCATCTCGAAATTCGTCCTCGGGGCA	584			
Db	77	GGACGAGCAATCATTTTCTTCTACGTATACCTACCCAGAACATTTCTTTAAACCAAGCA	136			
QY	585	CAGCTGGAAAGCTTTCGGAAACATCGTGCAGAACTACAGAGGCCGGCTGGTGAACCCAGG	644			
Db	137	CAACTCGAGATGCGAGAAAAATGTACATGATTACAGGCAAGAAATGTTTCCTCTCGT	196			
QY	645	ATCACCGAGGACCAAGCTGTGGAGGGCCAAAGTATGTGTAGTCCGCTTCATCCGAC	704			
Db	197	CTTACAGAAATGAAATGTGTGGAGCAAAAGTACATCTTGTGATTCAGCTTTTCATCTTAC	256			
QY	705	ACAGGGGAGAGTGGTCTGATTTGGCCGATGTCAGGCCAGTGCCTCATGAACATGACC	764			
Db	257	ACTGTGTGAAGATGATTTTGTATAGAAAGATGTTCAGGCCAGGTGCCATGAACATGACC	316			
QY	765	ATCACTGGCTCATGCTCACATTTCTACAGGAAGACCCCAACCGTGTGTTCTGCGAGTGG	824			
Db	317	ATCACAGGTTTATGATGATGACGTTTACAGGACTACGCCGCTGTCTGTTCTGCGCAGTGG	376			
QY	825	GTCGATAGTGCCTTCAATGCCATGTCTTAACTACATCCACCCAGTGGGACACTCCCATC	884			
Db	377	ATTAAACCAAGTCTTCAATGATCCGTGCTCAATTTACACCAACAGAAAGTGGAGAGCCACCCCTC	436			
QY	885	ACTGTGAGGACAGCTGGGGGACAGCCATGTGAGTGGCACACACTGAGAGCTGTGGCCAGGCC	944			
Db	437	ACTGTCAATGAGTTGGGAACAGCTTACTTTCTGTAAACAACGTGGGCCGTAGCAACAGCT	496			
QY	945	CTGGGACTCAAAATCCCTCACCAAGCACCTGCCCCCTTGGTGGCAGATTGTGCCCTTT	1004			
Db	497	CTAGGACATCATGCAATGTACCAAGCATGTCTACACACATGATAGGACGTTTGTTCCTCTTT	556			
QY	1005	GCAGAGTGGAGGCGCCCAATCGATCAAAATCCCCCGATAGGACAGAGAGG	1059			
Db	557	GCTGCGCGAGTGGCTGCTAAATTTGCAATTAATATTCATTAAATGAGCAAAAGCCATC	611			

```

RESULT 12
; US-10-206-901B-41
; Sequence 41, Application US/10206901B
; Publication No. US20030100540A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, ZHONGHUA
; APPLICANT: DUBOIS, RAYMOND
; TITLE OF INVENTION: IDENTIFICATION OF NSAID-REGULATED GENES
; FILE REFERENCE: VBLT:012US
; CURRENT APPLICATION NUMBER: US/10/206,901B
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/308,370
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 670
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-206-901B-41

Query Match      22.1%; Score 272.8; DB 14; Length 670;
Best Local Similarity 71.6%; Pred. No. 2,4e-72;
Matches 358; Conservative 0; Mismatches 142; Indels 0; Gaps 0.

Oy      733 CGATGTCAGCCAGGTGCCCATGAACATGACCATCATCAGGCTGTCATCATTTCTACA 792
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      4  GAATGTACGCCAGGTTCCCATGAAATGACCATACACAGTTGTATGATGACGTTTTCACA 63

Oy      793 GGAAGACCCCAACGATGGTGTCTTGCGCAGTGGGTGATCACTTCATAGCCATTGTTA 852

```

Db 64 GAGTACGGCGGCTGCTGCTTCTGCGACATGATTAACAGTCTTCAATGCGGTGTCA 123
QY 853 ACTATCCCAACCGCAGTGGTACACATCCCATCTGAGGCGGGGACGCTTATG 912
Db 124 ATTACACCAAGCAAGTGGAGAGCAGCCCTCTACTCTCATGTGTTGGAGAGCTTAGC 183
QY 913 TGAATCCCAACAGTGGAGTGTGGCCAGGCGCTGGAGCTCAATCTCCACCAAGCACC 972
Db 184 TTTCTGCAACAACTGTGGCCGTAGCAACAGCTCTAGACATCATGATGACCAACGATG 243
QY 973 TGGCCCCCTTGGCGGACATTTGTCCCTTTGACACATGGAGCTGCCAATGATCA 1032
Db 244 TCTCACCACTGATGAGACGCTTTTGTCTTGTCTGCTCCCTGCTGCTGCTGCTGCTGCT 303
QY 1033 ACATCCCGCTGATGAGAGAGAGAGCTCAGGTGGGATCCCGGCTGATGAGGACG 1092
Db 304 ATATTCCATTATGAGAGCAAGGAGACTCAAGTTGGCATTCCTCGTCAGGATGAGAAATG 363
QY 1093 GTACAGAGCTTGGCTACTCTGCTGACTGCAAGCCAAAGGGAATCTTCCAGTGTGATTT 1152
Db 364 GGAACCGCTTGGGGAGTGGCGGAGCTGCGCAACAGCATCAGCAAGTTGCTGTG 423
QY 1153 CAAGATCTGATGCGGATTTCTGCGATGGCCATCCCATGATGATGAGACATCTGG 1212
Db 424 CAGGATTTCTCATGGAGCGCCCTGGCATGGCCATCCCTCATTTATGATGACATTTGG 483
QY 1213 AGAAGAAAGACTTCTCTGAAG 1232
Db 484 AAAAGAAAGCTTTTGAAG 503

RESULT 13

US-09-918-995-23576
; Sequence 23576, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23576
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-23576

Query Match 17.2%; Score 212.4; DB 11; Length 489;
Best Local Similarity 69.6%; Pred. No. 4,3e-54;
Matches 288; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 446 AAGCATGAAACCAAAATGGTGAATTCCTTTAGACATCAACATCCAGAACTCGCTG 505
Db 76 AAGCTCCGGGACCATGCTGTAGAACTACCAACAACTTAACATCAAGAAACCTCGAG 135
QY 506 GGACCAAAAGTACTTCTTCTGGGAGAGCCGGCACTTTTACGCTGATGATCTCTGAAA 565
Db 136 GGAATCAAGCACTTCTTATGAGAGCAATCAATTTCTTACGTAAGTACCCAGGAA 195
QY 566 TGTGCTGCTGCTCGGGGACAGTGTGAACTTCTCGAAGATCTGTCAGAACTACAGGCG 625
Db 196 CATTCGTTAACCAAGCAACTGAGAGTGGGAGAAAATGATGATGATGACGCA 255
QY 626 CGCGCTGTGACCCAGGAGATCAGGAGACAGCTGTGAGGCGCAAGTATGTATGA 685
Db 256 AGGAATTTCTCTCTGCTTCTTACAGAAATGATGAGAGCAAGTACATCTATGTA 315

QY 686 CTCGCCCTTCATCCGAGACAGGAGGAGAGTGTCTGATTTGGCCGATGTACGCCCA 745
Db 316 TTCAGCTTTTCATCTGACAGCTGCTGGAAGATGATTTGATAGGAAGATGTACGCCCA 375
QY 746 GGTGCCATGAAATGACATCTACTGCTGATGCTCATTCTTACAGGAAGCCCAAC 805
Db 376 GGTTCCTCAAGACATGACATCAGAGTTGTATGATGACGCTTTTACAGGACTACGCCGC 435
QY 806 CGTGTGTTCTGAGCAGTGGGTGAATGATGCTTCAATGCCATGTTAACTACTC 859
Db 436 TGTGCTGTCTGCGACATGATTAACCAAGTCTTCAATGCCGTGCTCATTAACAC 489

RESULT 14

US-09-783-590-2110
; Sequence 2110, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillion, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16,261
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2110
; LENGTH: 293
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (35)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (82)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (159)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (172)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (177)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (181)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (212)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (225)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (233)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (262)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (275)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (281)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2003, 13:08:36 ; Search time 2505 Seconds

(without alignments)
11953.339 Million cell updates/sec

Title: us-09-990-415a-1

Perfect score: 1232

Sequence: 1 cccctagcgccgacgagacag.....agaagaagacttcctgaag 1232

Scoring table: IDENTITY, NUC

Searched: 22781392 seqs, 12152238056 residues 45562784

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database :

EST: *
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vit1:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	835.4	67.8	1201	13	BX376574 BX376574
2	824.6	66.9	914	9	AL521267 AL521267
3	816.4	66.3	869	13	BQ223306 AGENCOURT
4	798.8	64.8	998	12	BM473966 AGENCOURT

5	798.6	64.8	995	13	BX355800 BX355800
6	795.8	64.6	975	13	BQ933752 AGENCOURT
7	795.4	64.6	1201	13	BX334089 BX334089
8	773.8	62.8	777	12	B1908654 B1908654
9	767.2	62.3	840	9	AL522228 AL522228
10	759.6	61.7	811	12	B1818654 B1818654
11	759.4	61.6	1076	12	BM925231 BM925231
12	738.4	59.9	1014	13	BQ717821 BQ717821
13	727.2	57.0	862	13	BQ425697 BQ425697
14	713	57.9	963	13	BQ959114 BQ959114
15	711.6	57.3	1201	13	BX393485 BX393485
16	705.8	57.3	1040	12	BM548674 BM548674
17	700.4	56.9	961	12	BG775086 BG775086
18	700	56.8	755	12	B1517760 B1517760
19	691.2	56.1	888	12	BM018524 BM018524
20	690.8	56.1	1008	10	BF346247 BF346247
21	685.2	55.6	1008	10	BF569711 BF569711
22	680	55.2	1071	10	BF569784 BF569784
23	678.4	55.1	868	10	BE780804 BE780804
24	670.4	54.4	870	12	BM047107 BM047107
25	658.2	53.4	821	12	B1819328 B1819328
26	648.8	52.7	898	14	CB203758 CB203758
27	646.8	52.5	2682	11	AK034514 AK034514
28	617.8	50.1	634	9	AL701236 AL701236
29	614	49.8	934	13	B0506082 B0506082
30	583.2	47.3	764	14	CA317392 CA317392
31	581.6	47.2	731	14	CA317398 CA317398
32	577.6	46.9	746	12	B1752914 B1752914
33	576	46.8	715	14	CA316261 CA316261
34	562.8	45.7	1094	10	BG331175 BG331175
35	552.2	44.8	558	9	AL602426 AL602426
36	548	44.5	727	14	CD354095 CD354095
37	534.2	43.4	666	14	CA319631 CA319631
38	527.4	42.8	529	10	BE256336 BE256336
39	508.4	41.3	676	12	BM944910 BM944910
40	505	41.0	655	10	BB626574 BB626574
41	497.6	40.4	782	14	CB234292 CB234292
42	487.2	39.5	637	14	CA319259 CA319259
43	485.4	39.4	515	9	AM601313 AM601313
44	483.2	39.2	979	10	BG245508 BG245508
45	477.4	38.7	614	10	BB659747 BB659747

ALIGNMENTS

RESULT 1
LOCUS BX376574 1201 bp mRNA linear EST 08-MAY-2003
DEFINITION BX376574 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
ACCESSION BX376574
VERSION BX376574.1 GI:30448513
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORIGIN Homo sapiens

REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5549.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DD001DC03QPIc&cluster=5549.f. Contact :
Peng Liang Email: filiang@lifeotech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600

FEATURES	source	Location/Qualifiers
Faraday Avenue Genoscope sequence ID : CS0DD001DC03QP1.		
		1..1201
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="CS0DD001YF06"	
	/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"	
	/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"	
	/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."	
BASE COUNT		268 a 306 c 347 g 250 t 30 others
ORIGIN		

Query Match	67.8%	Score 835.4	DB 13	Length 1201
Best Local Similarity	98.5%	Pred. No. 4.6e-196		
Matches 842:	Conservative 1	Mismatches 12	Indels 0	Gaps 0
QY	378	GTGATGGCTGGAGAGGCCCGCGCGGACAGCGAGGACGAGAGAGAGGCGGTTCTGAGAG	437	
Db	41	GTACCGGTCGGAAATYTCGCCGGGATCGACAGCGGAGGACGAGAGGAAGCGGTTCTGAGAG	100	
QY	438	CTTAGAGAGCGGATGAGAAAGCAAAATGGGTGAATTGCCTTTCAGATCAATCAACGAGAA	497	
Db	101	CTTAGAGAGGAGATGAGAAAGCAAAATGGGTGAATTGCCTTTCAGATCAATCAACGAGAA	160	
QY	498	CCTGCGTGGGACCAAGACTTTCTCGGGACAGAGCCGGCACTTTTCAGGTTTCAGT	557	
Db	161	CCTGCGTGGGACCAAGACTTTCTCGGGACAGAGCCGGCACTTTTCAGGTTTCAGT	220	
QY	558	CCTGGAATCTGCTGCTGTCGGGGACAGCTGGAAGCTTCTCGGACATCGTGCAGAAC	617	
Db	221	CCTGGAATCTGCTGCTGTCGGGGACAGCTGGAAGCTTCTCGGACATCGTGCAGAAC	280	
QY	618	TACAGGGCGGGCGGTGGTACCCGAGGATCAACCGAGACAGCTGTGGAGGGCCAAATAT	677	
Db	281	TACAGGGCGGGCGGTGGTACCCGAGGATCAACCGAGACAGCTGTGGAGGGCCAAATAT	340	
QY	678	GTTGATGACTCGGCTTCATCCGGACACAGGGGAGAAAGTGGTCTGATTTGGCCGATG	737	
Db	341	GTTGATGACTCGGCTTCATCCGGACACAGGGGAGAAAGTGGTCTGATTTGGCCGATG	400	
QY	738	TCAGCCGAGGGGCCCATGAGACATATACATCTACTGGCTGATGCTCACATTTCTACAGAA	797	
Db	401	TCAGCCGAGGGGCCCATGAGACATATACATCTACTGGCTGATGCTCACATTTCTACAGAA	460	
QY	798	ACCCCAACCGTGTGTCTTGCGACAGTGGGTGAATAGTCCCTCAATGCATTGTAAACAC	857	
Db	461	ACCCCAACCGTGTGTCTTGCGACAGTGGGTGAATAGTCCCTCAATGCATTGTAAACAC	520	
QY	858	TCCAACCGGACGTGTGACACTCCCATCACTGTGAGGACAGTGGGGACAGCCTATGTGAGT	917	
Db	521	TCCAACCGGACGTGTGACACTCCCATCACTGTGAGGACAGTGGGGACAGCCTATGTGAGT	580	
QY	918	GCCACCACTGAGAGTGTGGCCAGGGCCCTGGGACATCAAAATCCCTACCAACACTCGCC	977	
Db	581	GCCACCACTGAGAGTGTGGCCAGGGCCCTGGGACATCAAAATCCCTACCAACACTCGCC	640	
QY	978	CCCTTGGTCGGCAGATTGTGCCCCCTTTTGACAGACAGTGGCAGCTGCCCACTGCATCAATC	1037	
Db	641	CCCTTGGTCGGCAGATTGTGTGCCCCCTTTTGACAGACAGTGGCAGCTGCCCACTGCATCAATC	700	
QY	1038	CCCTGTATGAGGACGAGAGAGCTCAGGTGGGCATCCGGTGGCTGATGAGGCAAGTCA	1097	
Db	701	CCCTGTATGAGGACGAGAGAGCTCAGGTGGGCATCCGGTGGCTGATGAGGCAAGTCA	760	
QY	1098	AGGCTTGGCTACTGCTGACAGCCAGAGGAGAAATCTTCAAGAGTGTATTTCAAGA	1157	
Db	761	AGGCTTGGCTACTGCTGACAGCCAGAGGAGAAATCTTCAAGAGTGTATTTCAAGA	820	
QY	1158	ATCTGCATGGCGATTTCTGCCAGTGGCATCCCACTGATCATGAGACACTGTGAGAG	1217	

[illegible]

```

Db      302 TGGCTGGGTTCTGCAATCCCGCTGCCCGGGTGGCGGGCGGGAAGCTCTCC 361
QY      343 CTCCTGCTGTCGGCGCTCAAGCGCTGACGTCCCGGATGATGGCTGGAGGGCGGGG 402
Db      362 CTCCTGCTGTCGGCGCTCAAGCGCTGACGTCCCGGATGATGGCTGGAGGGCGGGG 421
QY      403 GACAGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462
Db      422 GACAGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
QY      463 TGGGTGAATGCTTGTGACATCAATCCAGAGAACTCCGTGGAGCAAAAGTCTTCC 522
Db      482 TGGGTGAATGCTTGTGACATCAATCCAGAGAACTCCGTGGAGCAAAAGTCTTCC 541
QY      523 TGGGAGAGCGCGGCACTTTTCTACTGATCTCGAAATCTGCTGCTGTCGGGG 582
Db      542 TGGGAGAGCGCGGCACTTTTCTACTGATCTCGAAATCTGCTGCTGTCGGGG 600
QY      583 CACAGCTGGAAGCTTCTGGAACATCGTGAGAACTAGAGGCGCGGTGGTGAACCCAG 642
Db      601 CACAGCTGGAAGCTTCTGGAACATCGTGAGAACTAGAGGCGCGGTGGTGAACCCAG 660
QY      643 GGAATCACCGAGAGCACTGTGGAGGGCCAAAGTATGTATGATCCGCTTCATCCG 702
Db      661 GGAATCACCGAGAGCACTGTGGAGGGCCAAAGTATGTATGATCCGCTTCATCCG 720
QY      703 ACACAGGGGAGAGAGTGGTCTGATGGCGCATGTAGCCAGAGTCCCATGAACATGA 762
Db      721 ACACAGGGGAGAGAGTGGTCTGATGGCGCATGTAGCCAGAGTCCCATGAACATGA 780
QY      763 CCACTACTGCTGCACTCACTCACTTACAGAGAAAGCCCAACCGTGTCTTGAGCAGT 822
Db      781 CCACTACTGCTGCACTCACTCACTTACAGAGAAAGCCCAACCGTGTCTTGAGCAGT 839
QY      823 GGGGGAATCAGTCTTCAATGCCATGTGTAATCTCACTCCAGAGGAGTGAACCTCCA 882
Db      840 GGGGGAATCAGTCTTCAATGCCATGTGTAATCTCACTCCAGAGGAGTGAACCTCCA 899
QY      883 TCACGTGTAGAGCAGC 897
Db      900 TCACGTGTAGAGCAGC 914

RESULT 3
BQ223306 869 bp mRNA linear EST 02-MAY-2002
LOCUS AGENCOURT_7518440 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:606572
DEFINITION BQ223306.1 GI:20404706
ACCESSION BQ223306
VERSION BQ223306.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 869)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: egadps-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LAM13344 row: k column: 21
High quality sequence stop: 688.
location/Qualifiers
1. 869

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:606572"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_92"
/note="Organ: testis; Vector: PCMV-SPORE6; Site: 1; NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 193 a 244 c 255 g 175 t 2 others
ORIGIN
Query Match 66.3% Score 816.4; DB 13; Length 869;
Best Local Similarity 99.8%; Pred. No. 2e-191;
Matches 817; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
414 CAGAGAGAGAGCGGTTCTGAGAGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 473
Db 1 CAGAGAGAGAGCGGTTCTGAGAGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 474 CTTTACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 533
Db 61 CTTTACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 120
QY 534 CGGCACTTTTCACTGTTACTGATCGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 593
Db 121 CGGCACTTTTCACTGTTACTGATCGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 594 GCTTCTGGAACATCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 653
Db 181 GCTTCTGGAACATCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 654 GACCACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 713
Db 241 GACCACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 714 AAGGTGTCCTGATTTGGCCGCGCATGTCAAGCCAGAGTGCATGATGATGATGATGATGATG 773
Db 301 AAGGTGTCCTGATTTGGCCGCGCATGTCAAGCCAGAGTGCATGATGATGATGATGATG 360
QY 774 TGCATGCTCAATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 833
Db 361 TGCATGCTCAATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 834 TCCCTCAATGCCATTTTAACTACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 893
Db 421 TCCCTCAATGCCATTTTAACTACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 894 CAGCTGGGAGAGCGCTATGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 953
Db 481 CAGCTGGGAGAGCGCTATGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 954 AATTCCTCAACCAAGACCTGCCCCCTTGTGTGGAGAGATTTGTGCCCTTTGACAGAGTG 1013
Db 541 AATTCCTCAACCAAGACCTGCCCCCTTGTGTGGAGAGATTTGTGCCCTTTGACAGAGTG 600
QY 1014 GCAGCTGCCAATGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1073
Db 601 GCAGCTGCCAATGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 660
QY 1074 CCGGTGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1133
Db 661 CCGGTGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 1134 ATCTTCAGAGTGTGATTTCAAGAGATGCAATGAGAGAGATTCCTGCGCATGAGCATCCACA 1193
Db 721 ATCTTCAGAGTGTGATTTCAAGAGATGCAATGAGAGAGATTCCTGCGCATGAGCATCCACA 780
QY 1194 CTGATCATGAGACACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1232

```


BASE COUNT 226 a 281 c 281 g 193 t 14 others

ORIGIN

Query Match 64.8%; Score 798.6; DB 13; Length 995;
 Best Local Similarity 98.5%; Pred. No. 5.5e-187;
 Matches 795; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 426 CGGTTCTAGAGCTTCAGAGAGCCATGGAAGCAAAATGGGTAATTCCTTTAGACATC 485
 |||||||
 DB 67 CGGTTCTAGAGCTTCAGAGAGCCATGGAAGCAAAATGGGTAATTCCTTTAGACATC 126
 |||||||
 QY 486 AACATCCAGAAACCTCGGAGACCAAGTACTTCCGAGGACAGCCCGGACACTTTTC 545
 |||||||
 DB 127 AACATCCAGAAACCTCGGAGACCAAGTACTTCCGAGGACAGCCCGGACACTTTTC 186
 |||||||
 QY 546 ACTGTTACTGATTCCTCGAATCTGCTGCTCCGGGACACAGCTGGAAGCTTCTCGAAC 605
 |||||||
 DB 187 ACTGTTACTGAMCCCGGAATCTGCTGCTCCGGGACACAGCTGGAAGCTTCTCGAAC 246
 |||||||
 QY 606 ATCTGTCAGAACTACAGGCGCGGCTGCTGACCCAGGAGATCACCGAGACCACTGTGG 665
 |||||||
 DB 247 ATCTGTCAGAACTACAGGCGCGGCTGCTGACCCAGGAGATCACCGAGACCACTGTGG 306
 |||||||
 QY 666 AGGCGCAAGTATGTATGATCTCCGCTTCCATCCGACACAGGGGANAAGTGTCTG 725
 |||||||
 DB 307 AGGCGCAAGTATGTATGATCTCCGCTTCCATCCGACACAGGGGANAAGTGTCTG 366
 |||||||
 QY 726 ATGGCCGATGTCAGCCAGGAGTCCCATGAACATGACATCACTGCTGATGTCACA 785
 |||||||
 DB 367 ATGGCCGATGTCAGCCAGGAGTCCCATGAACATGACATCACTGCTGATGTCACA 426
 |||||||
 QY 786 TTTCACAGAAAGACCCCAACCGTGTGTTCTGGCAGTGGTGAATCACTCTTCATAGCC 845
 |||||||
 DB 427 TTTCACAGAAAGACCCCAACCGTGTGTTCTGGCAGTGGTGAATCACTCTTCATAGCC 486
 |||||||
 QY 846 ATTGTTAATCTACTCAACCCGAGTGTGACACTCCCATCACTGTGAGGACACTGGGACA 905
 |||||||
 DB 487 ATTGTTAATCTACTCAACCCGAGTGTGACACTCCCATCACTGTGAGGACACTGGGACA 546
 |||||||
 QY 906 GCCATATGAGTGCACACACTGAGAGCTGTGGCCAGGCGCTGGGACTCAAAATCCCTCAC 965
 |||||||
 DB 547 GCCATATGAGTGCACACACTGAGAGCTGTGGCCAGGCGCTGGGACTCAAAATCCCTCAC 606
 |||||||
 QY 966 AAGACCTGCCCCCTTGTGTGGGAGATTTGTGCCCTTTGCAGCAGTGGCCACAC 1025
 |||||||
 DB 607 AAGACCTGCCCCCTTGTGTGGGAGATTTGTGCCCTTTGCAGCAGTGGCCACAC 666
 |||||||
 QY 1026 TGCATCAACATCCCCCTGATGAGGACAGAGAGCTGAGTGGGATCCGCTGGCTGAT 1085
 |||||||
 DB 667 TGCATCAACATCCCCCTGATGAGGACAGAGAGCTGAGTGGGATCCGCTGGCTGAT 726
 |||||||
 QY 1086 GAGCAGGTCAGAGGCTTGGCTGACTCGGTGACTGACAGCAAGCAGGAAATCTTCAGGTTG 1145
 |||||||
 DB 727 GAGCAGGTCAGAGGCTTGGCTGACTCGGTGACTGACAGCAAGCAGGAAATCTTCAGGTTG 786
 |||||||
 QY 1146 GTGATTTCAAGAATCTGATGGGATTCCTGCAATCCCATGCGCATCCACACTATATGAGAC 1205
 |||||||
 DB 787 GTGATTTCAAGAATCTGATGGGATTCCTGCAATCCCATGCGCATCCACACTATATGAGAC 846
 |||||||
 QY 1206 ACTCTGAGAGAAAGACTTCTCTGAG 1232
 |||||||
 DB 847 ACTCTGAGAGAAAGACTTCTCTGAG 873
 |||||||

RESULT 6
 BQ933752 975 bp mRNA linear EST 21-AUG-2002
 LOCUS AGENCOURT.8837545 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6424845
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ933752.1 GI:22349135
 VERSION BQ933752.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 975)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DPV/Gazdar
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULNL at:
<http://image.llnl.gov>
 Plate: L1CM2605 row: k column: 22
 High quality sequence stop: 646.
 Location/Qualifiers
 1..975
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6424845"
 /tissue_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH-MGC_18"
 /note="Organ: lung; Vector: pORF7; Site:1: XhoI; Site:2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH-MGC Library."

BASE COUNT 225 a 277 c 267 g 199 t 7 others

ORIGIN

Query Match 64.6%; Score 795.8; DB 13; Length 975;
 Best Local Similarity 99.1%; Pred. No. 2.7e-186;
 Matches 800; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 421 GAAGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTAATTCCTTTAG 480
 |||||||
 DB 1 GAAGCGGTTCTGAGAGCTTCAGAGAGCGATGGAAGCAAAATGGGTAATTCCTTTAG 60
 |||||||
 QY 481 ACATCAACATCCAGAAACCTGCTGGGACCAAAAGTACTTCTCGGAGAGCCGGCACT 540
 |||||||
 DB 61 ACATCAACATCCAGAAACCTGCTGGGACCAAAAGTACTTCTCGGAGAGCCGGCACT 120
 |||||||
 QY 541 TTTTCACTGTACTGATCCTGAAATCTGCTGTCTCGGGGACACAGCTGGAAGCTTCTC 600
 |||||||
 DB 121 TTTTCACTGTACTGATCCTGAAATCTGCTGTCTCGGGGACACAGCTGGAAGCTTCTC 180
 |||||||
 QY 601 GGAACATCTGCAAACTACAGGCGCGGCTGTGACCCAGGATCAACAGACCAAGC 660
 |||||||
 DB 181 GGAACATCTGCAAACTACAGGCGCGGCTGTGACCCAGGATCAACAGACCAAGC 240
 |||||||
 QY 661 TGTGAGAGGCCAATATGTATGATCTCCGCTTCCATCCGAGACACAGGAGGAAGTGG 720
 |||||||
 DB 241 TGTGAGAGGCCAATATGTATGATCTCCGCTTCCATCCGAGACACAGGAGGAAGTGG 300
 |||||||
 QY 721 TCTGATTTGGCGCATGTGACGCCAGGTGCCATGAACTACATCACTGCTGCTGATGC 780
 |||||||
 DB 301 TCTGATTTGGCGCATGTGACGCCAGGTGCCATGAACTACATCACTGCTGCTGATGC 360
 |||||||
 QY 781 TCACATTTCTACAGGAAGACCCCAACCGTGTGCTGAGTGGTGAATGATGCTCTCA 840
 |||||||
 DB 361 TCACATTTCTACAGGAAGACCCCAACCGTGTGCTGAGTGGTGAATGATGCTCTCA 420
 |||||||
 QY 841 ATGCCATTTGTTAACTACTCCAAACGCAAGTGTGACACTCCATCACTGTGAGGACAGCTG 900
 |||||||
 DB 421 ATGCCATTTGTTAACTACTCCAAACGCAAGTGTGACACTCCATCACTGTGAGGACAGCTG 480
 |||||||

Query Match	64.6%	Score 795.4	DB 13	Length 1201
Best Local Similarity	99.3%	Pred. No. 3.7e-166		
Matches 807; Conservative	2;	Mismatches 3;	Indels 1;	Gaps 1;

RESULT 8
 BI908654
 LOCUS
 DEFINITION
 BI908654 777 bp mRNA linear EST 16-OCT-2001
 60306677.F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5215818 5',
 mRNA sequence.
 ACCESSION
 VERSION BI908654
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 777)
 NIH-MGC <http://mgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 JOURNAL

Db	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	154	CCCTCCGGCTCGAACCACTCTCTCTGAGCCAGGACAGTACCGGGGCTCTGGAAATTG						
QY	186	CCACCCCTCCCTGGGACACCTTGAGAGCCCTTCCTGAGAGGACGTACACGGGGCAGAGCGGGA						
Db	214	CCACCCCTCCCTGGGACACCTTGAGAGCCCTTCCTGAGAGGACGTACACGGGGCAGAGCGGGA						
QY	246	CGTAGACCTGAGTTTGCTCAGAGCGGCTCTGTTGTTGGTGGGTTTGCCAAATCCCGG						
Db	274	CGTAGACCTGAGTTTGCTCAGAGCGGCTCTGTTGTTGGTGGGTTTGCCAAATCCCGG						
QY	306	TGCGCACCGGGTGGGCGCGCGGGAAGTCTCTGCCCTCCCTGCTGTGGCGCTCAGC						
Db	334	TGCGCACCGGGTGGGCGCGCGGGAAGTCTCTGCCCTCCCTGCTGTGGCGCTCAGC						
QY	366	GTGACGTCCCGGGTGTGGCTGGAGGGCCCGCGCGACACCGAGAGCAGAGAGAGG						
Db	394	GTGACGTCCCGGGTGTGGCTGGAGGGCCCGCGCGACACCGAGAGCAGAGAGAGG						
QY	426	CGGTTCTGAGACTTCAGAGACGATGGAACCAAAATGGGTGAATTGCTTTAGACATC						
Db	454	CGGTTCTGAGACTTCAGAGACGATGGAACCAAAATGGGTGAATTGCTTTAGACATC						
QY	486	AACATCCAGAGAACCTTCGCTGGGACCAAACTCTTCTTGAGAGACCGCGACCTTTTC						
Db	514	AACATCCAGAGAACCTTCGCTGGGACCAAACTCTTCTTGAGAGACCGCGACCTTTTC						
QY	546	ACTGTTACGATCTCTGAAATCTGCTGCTCCGGGGGACACACTGGAAGCTTCTGGAGAC						
Db	573	ACTGTTACGATCTCTGAAATCTGCTGCTCCGGGGGACACACTGGAAGCTTCTGGAGAC						
QY	606	ATCGTGCAGAACTACAGGGCGCGGCTGGTGACCCAGGAGATCACAGAGACCAAGCTGTGG						
Db	633	ATCGTGCAGAACTACAGGGCGCGGCTGGTGACCCAGGAGATCACAGAGACCAAGCTGTGG						
QY	666	AGGGCCAGTAGTGTATGATGATCGGCGCTTCATCCGGACACAGGGGAGAAAGTGCTCTG						
Db	693	AGGGCCAGTAGTGTATGATGATCGGCGCTTCATCCGGACACAGGGGAGAAAGTGCTCTG						
QY	726	ATTGGCCGCAATGTGACGCCAGGTGCCATGAAATGACATCACTAGCTGATGCTGCACA						
Db	753	ATTGGCCGCAATGTGACGCCAGGTGCCATGAAATGACATCACTAGCTGATGCTGCACA						
QY	786	TTCTACAGAGAGACCCCAACCGTGTGT 813						
Db	813	TTCTACAGAGAGACCCCAACCGTGTGT 840						
RESULT 10								
Bi18854	811 bp	mRNA	linear	EST 04-OCT-2001				
LOCUS	603037323p1	NIH_MGC_115	Homo sapiens cDNA clone IMAGE:5178437 5',					
DEFINITION			mRNA sequence.					
ACCESSION	Bi18854							
KEYWORDS	Bi18854.1	GI:15930404						
SOURCE								
ORGANISM			Homo sapiens (human)					
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
JOURNAL			NIH-MGC http://mhc.nci.nih.gov/.					
COMMENT			National Institutes of Health, Mammalian Gene Collection (MGC)					
			Unpublished					
			Contact: Robert Strausberg, Ph.D.					
			Email: cgabbs@mail.nih.gov					
			Tissue Procurement: Life Technologies, Inc.					
			cDNA Library Preparation: Life Technologies, Inc.					
			DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)					
			DNA Sequencing by: Incyte Genomics, Inc.					
			Clone distribution: MGC clone distribution information can be					
			found through the I.M.A.G.E. Consortium/LNL at:					
			http://image.lnl.gov					

```

Plate: ILWMI1444 row: n column: 06
High quality sequence stop: 784.
Location/Qualifiers
1. 811
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5178437"
/lab_host="DH10B"
/clone_11b="NIR_MCC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-OT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIR_MCC Library."
BASE COUNT
180 a 223 c 246 g 162 t
ORIGIN
```

Query Match	Similarity	61.7%	Score 759.6	DB 12	Length 811
Best Local	Similarity	99.0%	Pred. Mismatch 17		
Matches	806	Conservative	0	Indels	4
					Gaps 4
QY	378	GTGATGGCTGGAGAGGCCCGGCGGCGAGACGCCGAGAGCAGAGAGAGGAGCGGTTCTGAGAG	437		
Db	1	GTGATGGCTGGAGAGGCCCGGCGGCGAGAGCAGCGGAGAGAGAGAGGAGCGGTTCTGAGAG	60		
QY	438	CTTCAGAGAGCGATGGAAAGCAAAATGGGTGAATTGCCTTTAGACATCAACATCCAGAA	497		
Db	61	CTTCAGAGAGCGATGGAAAGCAAAATGGGTGAATGCGCTTTAGACATCAACATCCAGAA	119		
QY	498	CCTCGCTGGAGCAAAAGTACTTTCCTGGGACAGAGCCGCGCACTTTTACGTGTACGTAT	557		
Db	120	CCTCGCTGGAGCAAAAGTACTTTCCTGGGACAGAGCCGCGCACTTTTACGTGTATCTAT	178		
QY	558	CCTGAAATCTGCTGCTGTCCGGGGACACAGCTGGAGGCTTCCGAAATCGTCAGAAC	617		
Db	179	CCTGAAATCTGCTGCTGTCCGGGGACACAGCTGGAGGCTTCCGAAATCGTCAGAAC	238		
QY	618	TACAGAGCGCGGGGTGAGACCCCAAGGGATCAACCGAGGACAGCTGTGGAGGGCCAAAGT	677		
Db	239	TACAGAGCGCGGGGTGAGATACCCCAAGGGATCAACCGAGGACAGCTGTGGAGGGCCAAAGT	298		
QY	678	GTGATGACTCCGCTTCATCCGAGACACAGAGGGAGGAAGTGTCGTGATTGGCCGATG	737		
Db	299	GTGATGACTCCGCTTCATCCGAGACACAGAGGGAGGAAGTGTCGTGATTGGCCGATG	358		
QY	738	TCAGGCCAGGTGCCCATGAAATACATACCATCTGCTGATGCTTCACATTTCTACAGAG	797		
Db	359	TCAGGCCAGGTGCCCATGAAATACATACCATCTGCTGATGCTTCACATTTCTACAGAG	418		
QY	798	ACCCCAACCGTGTCTCTGCGACAGTGGGTGAATAGTCCCTCAATGGCATTTGTAACATC	857		
Db	419	ACCCCAACCGTGTCTCTGCGACAGTGGGTGAATAGTCCCTCAATGGCATTTGTAACATC	478		
QY	858	TCCAACCGCAGTGTGACATCTCCATCATCTGTGAGGAGCAGTGGGAGACGCTTATGTAGT	917		
Db	479	TCCAACCGCAGTGTGACATCTCCATCATCTGTGAGGAGCAGTGGGAGACGCTTATGTAGT	538		
QY	918	GCCACCACTGGAGTGTGGCCACAGGCGCTGGGACATCAAAATCCCTACACAGCAACCTGCC	977		
Db	539	GCCACCACTGGAGTGTGGCCACAGGCGCTGGGACATCAAAATCCCTACACAGCAACCTGCC	598		
QY	978	CCCTTGGTCGGCAGATTGTGCCCCCTTTCAGACAGTGGCAGTGGCACTGCATCAACATC	1037		
Db	599	CCCTTGGTCGGCAGATTGTGCCCCCTTTCAGACAGTGGCAGTGGCACTGCATCAACATC	658		
QY	1038	CCCTTGTATGAGGACAGAGAGCTGACAGGTGGGACATCCCGTGGCTGATGAGCAGTCAAG	1097		

Db 659 CCCCTGATGAGCAGAGAGACTGCAGAGTGGGATCCCGGTGCTGATGAGCGAGTTCAG 718

QY 1098 AGCCTGGCTACTCGGTGACTGCAGCCAG-CAGGAAATCTTCAGAGTGGTGAATTCAG 1156

Db 719 AGCCTGGCTACTCGGTGACTGCAGAGAGAGGAAATCTTCAGAGTGGTGAATTCAG 778

QY 1157 AATCTGCATGGCGATTCCTGCATGGCGATCCCA 1190

Db 779 AATCTGCATGGCGA-TCTGCATGGCGATCCCA 811

RESULT 11
BM925231 1076 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT_6627462 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5762556
DEFINITION 5', mRNA sequence.
ACCESSION BM925231
VERSION BM925231.1 GI:19375610
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1076)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LRAM12812 row: p column: 13
High quality sequence start: 4
High quality sequence stop: 664.
Location/Qualifiers
1. 1076
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5762556"
/lab_host="DH10B"
/clone_1bp="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is c190-dt primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb. Insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library."

BASE COUNT 230 a 316 c 314 g 216 t

ORIGIN

Query Match 61.6%; Score 759.4; DB 12; Length 1076;
Best Local Similarity 99.7%; Pval. No. 2.9e-177;
Matches 771; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 460 AATGGGTGAATTCCTTAGACATCAATCCAGAACTCGTGGACCAAGTACTT 519

Db 11 AATGGGTGAATTCCTTAGACATC-ACATCCAGAACTCGTGGACCAAGTACTT 69

QY 520 TCCCTGGCAGAGCCCGCACTTTTCACTGTACTGATCCGAAATCTGCTGTCG 579

Db 70 TCCCTGGCAGAGCCCGCACTTTTCACTGTACTGATCCGAAATCTGCTGTCG 129

QY 580 GGGCAGAGCTGGAAGCTTTCGCGAATCGTGCAGAACTACAGGCGCGCTGTGACCC 639

Db 130 GGGCAGAGCTGGAAGCTTTCGGAACATCGTGCAGAACTACAGGCGCGGTGTGACCC 189

QY 640 CAGAGATCACAGAGAGACACTGTGGAGGCGCAATGATGATAGATCCGCTTCATC 699

Db 190 CAGGATCACAGAGAGACACTGTGGAGGCGCAATGATGATAGATCCGCTTCATC 249

QY 700 CGGACACAGGAGAGAGTGGTCTGATTTGGCCCATGTACGCCAGGTGCCATGACA 759

Db 250 CGGACACAGGAGAGAGTGGTCTGATTTGGCCCATGTACGCCAGGTGCCATGACA 309

QY 760 TGACCATCACTGGCTGATGCTCATTTACAGAGAGACCCCAAGCTGGTGTTCGGC 819

Db 310 TGACCATCACTGGCTGATGCTCATTTACAGAGAGACCCCAAGCTGGTGTTCGGC 369

QY 820 AGTGGGTGAATCACTGCTTCATATGCAATGTTTAACTACTCCACCGAGTGGTACACTC 879

Db 370 AGTGGGTGAATCACTGCTTCATATGCAATGTTTAACTACTCCACCGAGTGGTACACTC 429

QY 880 CCATCACTGTGAGGACAGCTGGGACAGACCTATGTAGTCCACCACTGAGCTGTGCCA 939

Db 430 CCATCACTGTGAGGACAGCTGGGACAGACCTATGTAGTCCACCACTGAGCTGTGCCA 489

QY 940 CGGCGCTGGGACATCAATTCCTTCCACCAAGACCTGCCCCCTTGTGGCGATTTGTC 999

Db 490 CGGCGCTGGGACATCAATTCCTTCCACCAAGACCTGCCCCCTTGTGGCGATTTGTC 549

QY 1000 CTTTGCAGAGTGGACAGCTGCGCAACCTGCATCAATCCCTATGAGGACAGAGAGC 1059

Db 550 CTTTGCAGAGTGGACAGCTGCGCAACCTGCATCAATCCCTATGAGGACAGAGAGC 609

QY 1060 TGCAGTGGGACATCCCGGTGGCTGATGAGGACAGCTGAGAGCTTGGCTACTGCTG 1119

Db 610 TGCAGTGGGACATCCCGGTGGCTGATGAGGACAGCTGAGAGCTTGGCTACTGCTG 669

QY 1120 CAGCAGAGCAGGAGATCTTCAGGTGGATTTGAAGATTCGATGAGCGATTCGTGCCA 1179

Db 670 CAGCAGAGCAGGAGATCTTCAGGTGGATTTGAAGATTCGATGAGCGATTCGTGCCA 729

QY 1180 TGGCCATCCCAACCATGATGACACTGTGGAAGAAAGACTTCTGTGAAG 1232

Db 730 TGGCCATCCCAACCATGATGACACTGTGGAAGAAAGACTTCTGTGAAG 782

RESULT 12
BQ717821 1014 bp mRNA linear EST 16-JUL-2002
LOCUS AGENCOURT_8229525 LupsK1.dorsal.root.ganglion Homo sapiens cDNA
DEFINITION clone IMAGE:6184833 5', mRNA sequence.
ACCESSION BQ717821
VERSION BQ717821.1 GI:21856718
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1014)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: Dr. James R. LupsK1
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LRAM13574 row: k column: 10
High quality sequence stop: 610.
Location/Qualifiers
1. 1014

FEATURES
source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6184833"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski.dorsal_root_ganglion"
/note="vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGCCG-3' and
5'-GACATCTCTAGATCGCGAGCGGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

```

BASE COUNT 199 a 292 c 317 g 206 t

ORIGIN

Query Match 59.9%; Score 738.4; DB 13; Length 1014;
 Best Local Similarity 96.5%; Pred. No. 4.5e-172;
 Matches 799; Conservative 0; Mismatches 21; Indels 8; Gaps 4;

```

QY 222 GGAGCTGACGGGCGAGAGCGGAGCTGAGCTGAGTTGGTGCAGGGGCTGCTGTG 281
DB 22 GGAATGTCACGGGCGAGAGCGGAGCTGAGCTGAGTTGGTGCAGGGGCTGCTGTG 81
QY 282 GTGGCTGGGTTCTGCAATCCCGTCCACCGGGTGGGCGCGCGGGAAGCTCTGCC 341
DB 82 GTGGCTGGGTTCTGCAATCCCGTCCACCGGGTGGGCGCGCGGGAAGCTCTGCC 141
QY 342 CCTCCCTGCTGGTGGGCGTACGGGTGACGTCCTCCGCGTGAAGGCTGGGAGGCGCGG 401
DB 142 CCTCCCTGCTGGTGGGCGTACGGGTGACGTCCTCCGCGTGAAGGCTGGGAGGCGCGG 201
QY 402 CGACAGGCGAGGCGAGAGAGGCGGTTCTGAGAGCTTCAGAGAGCATGGAAGCAA 461
DB 202 CGACAGGCGAGGCGAGAGAGGCGGTTCTGAGAGCTTCAGAGAGCATGGAAGCAA 261
QY 462 ATGGGTAATTCCTTTAGACATCAACATCCAGAACCTCGCTGGGACCAAGTACTTTC 521
DB 262 ATGGGTAATTCCTTTAGACATCAACATCCAGAACCTCGCTGGGACCAAGTACTTTC 321
QY 522 CTGGGCGAGAGCGCGCACTTTTACAGTCTGATCTGTAATCTGCTGCTGCTGCGGG 581
DB 322 CTGGGCGAGAGCGCGCACTTTTACAGTCTGATCTGTAATCTGCTGCTGCTGCGGG 381
QY 582 GCACAGCTGGAAGCTTCTCGGAACATCGTGAAGACTAGAGAGGCGCGGCTGTACCCCA 641
DB 382 GCACAGCTGGAAGCTTCTCGGAACATCGTGAAGACTAGAGAGGCGCGGCTGTACCCCA 441
QY 642 GGGATCCAGGAGGCGAGCTGTGAGAGGCGCAAGTATGTATGACTCCGCTTCATCCG 701
DB 442 GGGATCCAGGAGGCGAGCTGTGAGAGGCGCAAGTATGTATGACTCCGCTTCATCCG 501
QY 702 GACACAGGAGGAGAGGTGCTGATGAGCGGATGTCAGCGCCAGGAGCCCATGAATG 761
DB 502 GACACAGGAGGAGAGGTGCTGATGAGCGGATGTCAGCGCCAGGAGCCCATGAATG 561
QY 762 ACCATCACTGCTGATGCTACATTTCTACAGGAAGACCCCAACCTGTGTTCTGCGAG 821
DB 562 ACCATCACTGCTGATGCTACATTTCTACAGGAAGACCCCAACCTGTGTTCTGCGAG 621
QY 822 TGGGTGAATGATCCCTCAATGCAATGTTAACTACTCAACGCAAGTGTGACATGCC 881
DB 622 TGGGTGAATGATCCCTCAATGCAATGTTAACTACTCAACGCAAGTGTGACATGCC 681
QY 882 ATCACTGTGAGGAGCTGGGAGACGCTATGTAGTCCACCACTGAGGCTGTGGCCAG 941
DB 682 ATCACTGTGAGGAGCTGGGAGACGCTATGTAGTCCACCACTGAGGCTGTGGCCAG 741

```

```

QY 942 GCCCTGGGAGCTC-AAATCCCTTACCAAGACCTGCCCC---TTGGTGGAGATT--- 994
DB 742 GCCCTGGGAGCTC-AAATCCCTTACCAAGACCTGCCCC---TTGGTGGAGATTGG 801
QY 995 TGTGGCCTTTGACGAGCTGGGAGCTG-CCAACTGCAATCAATCCCC 1041
DB 802 TGTGGCCTTTGACGAGCTGGGAGCTGCAAACTGCAATCAATCCCC 849

```

RESULT 13
 B0425697
 LOCUS
 DEFINITION
 AGENCOURT_7887371 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6166944
 B0425697
 B0425697.1 GI:21121012
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC/DCPD/DRP
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M13528 row: b column: 01
 High quality sequence stop: 603.
 Location/Qualifiers
 1. 862
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6166944"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies"

FEATURES

BASE COUNT 200 a 249 c 246 g 166 t 1 others

ORIGIN

Query Match 59.0%; Score 727.2; DB 13; Length 862;
 Best Local Similarity 98.7%; Pred. No. 2.5e-169;
 Matches 743; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

```

QY 395 CCGGCGGCGAGACGCGGAGGCGAGAGAGGCGGTTCTGAGAGCTTCAGAGCGAGTGA 454
DB 1 CCGGCGGCGAGACGCGGAGGCGAGAGAGGCGGTTCTGAGAGCTTCAGAGCGAGTGA 60
QY 455 AAGCAAAATGGTGAATTCCTTTAGACATCAATCAGGAACCTGCGGGGACCAAG 514
DB 61 AAGCAAAATGGTGAATTCCTTTAGACATCAATCAGGAACCTGCGGGGACCAAG 120
QY 515 TACTTTCCTGGGAGAGCCCGGACCTTTTACATGTTACTGATCCTGAAATCTGCTGCT 574
DB 121 TACTTTCCTGGGAGAGCCCGGACCTTTTACATGTTACTGATCCTGAAATCTGCTGCT 180
QY 575 GTCCGGGCGACACGCTGGAACCTTCTCGAATCTGTCAGAACTACAGGCGCGGCTGCT 634
DB 181 GTCCGGGCGACACGCTGGAACCTTCTCGAATCTGTCAGAACTACAGGCGCGGCTGCT 240

```

635 GACCCAGGATCAACCCAGACAGCTGTGAGAGGCCAAGTATGTATGATCCGCTT 694
695 CCATCCGACACAGAGGAGAGTGTCTGATTTGGCCGATGTCAGCCAGTGGCCAT 754
301 CCATCCGACACAGGAGGAGAGTGTCTGATTTGGCCGATGTCAGCCAGTGGCCAT 360
755 GAACATGACATCACTGGCTGCATGCTCACAATTTCTACAGAAACCCCAACCGTGTG 814
361 GAACATGACATCACTGGCTGCATGCTCACAATTTCTACAGAAACCCCAACCGTGTG 420
815 CTGCAAGTGGTGAATGATGCTCTCAATGCTTCAATGCTTCAATGCTTCAATGCTTCA 874
421 CTGCAAGTGGTGAATGATGCTCTCAATGCTTCAATGCTTCAATGCTTCAATGCTTCA 480
875 CACTCCATCACTGTGAGGAGCAGTGGGAGACAGCTATGTAGTGCA-CCACTGAGCTG 933
481 CACTCCATCACTGTGAGGAGCAGTGGGAGACAGCTATGTAGTGCA-CCACTGAGCTG 540
934 TGGCCACGCGCCCTGGGACTCAATTCCTCACCAGACACCTGCCCTTGGTGGCAGAT 993
541 TGGCCACGCGCCCTGGGACTCAATTCCTCACCAGACACCTGCCCTTGGTGGCAGAT 600
994 TTTGGCCCTTGGAGCAGTGGGACTGCTCCAACTGATCAATTCCTTGTGAGGAGCA 1053
601 TTTGGCCCTTGGAGCAGTGGGACTGCTCCAACTGATCAATTCCTTGTGAGGAGCA 660
1054 GAGAGCTGACAGTGGGAGTCCCGGTGGCTGATGAGGAGCAGTGGGCTTGGTGGTGG 1113
661 GAGAGCTGACAGTGGGAGTCCCGGTGGCTGATGAGGAGCAGTGGGCTTGGTGGTGG 720
1114 TGAATGACAGCCAGAGGAAATCTTCCAGTGG 1146
721 TGAATGACAGCCAGAGGAAATCTTCCAGTGG 753

RESULT 14
LOCUS B0959114 963 bp mRNA linear EST 21-AUG-2002
DEFINITION AGNCOURT_10031240 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6482227
5', mRNA sequence.
B0959114
ACCESSION B0959114.1 GI:22374592
VERSION B0959114.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 963)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgsb@remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LNCM2664 row: n column: 20
High quality sequence stop: 635.
Location/Qualifiers
1. 963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6482227"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_40"

FEATURES
source

/note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI;
Site: 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: This is a NIH-MGC Library."

BASE COUNT 200 a 290 c 272 g 201 t

ORIGIN

Query Match 57.9%; Score 713; DB 13; Length 963;
Best Local Similarity 100.0%; Pred. No. 8.6e-166;
Matches 713; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

520 TCCTGGGAGAGACCCGGGACATTTTCACTGATGATCTGCAATCTGCTGCTCG 579
1 TCCTGGGAGAGACCCGGGACATTTTCACTGATGATCTGCAATCTGCTGCTCG 60
580 GGGCAGACAGTGAAGCTTCTCGAATCATGTCAGAGACAGGCGCGGTGTGACCC 639
61 GGGCAGACAGTGAAGCTTCTCGAATCATGTCAGAGACAGGCGCGGTGTGACCC 120
640 CAGGATACACGAGAGACAGCTGTGAGAGGCGCAATATGTATGATCTCCGCTTCCATC 699
121 CAGGATACACGAGAGACAGCTGTGAGAGGCGCAATATGTATGATCTCCGCTTCCATC 180
700 CGGACACAGGAGAGAGGAGTGTCTGATTTGGCCGATGTACAGCCAGTGCATGAGACA 759
181 CGGACACAGGAGAGAGGAGTGTCTGATTTGGCCGATGTACAGCCAGTGCATGAGACA 240
760 TGACCATCACTGGCTGATGCTCAATTCCTCAGAGAGAGAGAGAGAGAGAGAGAG 819
241 TGACCATCACTGGCTGATGCTCAATTCCTCAGAGAGAGAGAGAGAGAGAGAGAG 300
820 AGTGGGGAATCAATCTCTTCAATGCCATGTTAACTACTCCAGCCGAGTGTGACTC 879
301 AGTGGGGAATCAATCTCTTCAATGCCATGTTAACTACTCCAGCCGAGTGTGACTC 360
880 CCATCACTGTGAGGAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 939
361 CCATCACTGTGAGGAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
940 CGGCGCTGGGAGCTCAATTCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 999
421 CGGCGCTGGGAGCTCAATTCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
1000 CCTTTGAGAGAGTGGGAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1059
481 CCTTTGAGAGAGTGGGAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
1060 TGCAGTGGGAGATCCCGGTGGCTGATGAGGAGAGAGAGAGAGAGAGAGAGAG 1119
541 TGCAGTGGGAGATCCCGGTGGCTGATGAGGAGAGAGAGAGAGAGAGAGAGAG 600
1120 CAGCCAGAGAGAGAGATCTTCCAGGTGTGATTTCAAGATTCGATGGCGATTCTGCA 1179
601 CAGCCAGAGAGAGAGATCTTCCAGGTGTGATTTCAAGATTCGATGGCGATTCTGCA 660
1180 TGGCCATCCCAACCATGATCATGACACTGGGAGAGAGAGAGAGAGAGAGAGAG 1232
661 TGGCCATCCCAACCATGATCATGACACTGGGAGAGAGAGAGAGAGAGAGAGAG 713

RESULT 15
LOCUS BX393485 1201 bp mRNA linear EST 13-MAY-2003
DEFINITION BX393485 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
CDNA clone CS0DB004YM02 5-PRIME, mRNA sequence.
ACCESSION BX393485
VERSION BX393485.1 GI:30616089
KEYWORDS EST.
SOURCE Homo sapiens (human)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 16:06:27 ; Search time 261 Seconds

(without alignments)
2626.128 Million cell updates/sec

Title: US-09-990-415A-2

Perfect score: 1345

Sequence: 1 MESKXGELPLDINIQEPKWD.....PAMAIPLIMDTLEKKDFLK 261

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n_model -DEV=xlp
-O=/cgn2_1/USPTO_spool_P/US09990415/rnat_16102003_073808_9492/app_query.fasta_1.455
-DB=Published Applications NA -PMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bits -SHARP=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09990415.@CGN_1.1.85_rnat_16102003_073808_9492
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-GAPOP=6 -GAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA.*
1: /cgn2_6/ptodata/2/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubna/PCI_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubna/US06_NEM_PUB.seq.*
4: /cgn2_6/ptodata/2/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubna/US07_NEM_PUB.seq.*
6: /cgn2_6/ptodata/2/pubna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubna/US08_NEM_PUB.seq.*
8: /cgn2_6/ptodata/2/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubna/US09A_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1345	100.0	1232	US-09-990-415A-1	Sequence 1, Appl

2	1345	100.0	1375	9	US-09-867-550-1343	Sequence 1343, Ap
3	1345	100.0	1729	14	US-10-037-270-1016	Sequence 1016, Ap
4	1089	81.0	1134	14	US-10-014-338-1	Sequence 1, Appl
5	1089	81.0	2269	10	US-09-990-415A-7	Sequence 7, Appl
6	1089	81.0	2878	14	US-10-198-846-9767	Sequence 9767, Ap
7	1003.5	74.6	1061	10	US-09-990-415A-3	Sequence 3, Appl1
8	897.5	66.7	984	14	US-10-014-338-3	Sequence 3, Appl1
9	754	56.1	1567	10	US-09-990-415A-5	Sequence 5, Appl1
10	750	55.8	2376	13	US-10-098-841-64	Sequence 64, Appl
11	750	55.8	2458	9	US-09-822-849A-40	Sequence 490, Appl
12	708	52.6	670	14	US-10-206-901B-41	Sequence 41, Appl
13	587	43.6	489	11	US-09-918-995-23576	Sequence 23576, A
14	447.5	33.3	3342	12	US-09-814-353-20153	Sequence 20153, A
15	358	26.6	522	11	US-09-918-995-5675	Sequence 5675, Ap
16	344	25.6	293	10	US-09-783-590-2110	Sequence 2110, Ap
17	321	23.9	648	14	US-10-198-846-7595	Sequence 7595, Ap
18	295	21.9	533	11	US-09-918-995-26900	Sequence 26900, A
19	230	17.1	238	11	US-09-919-936-1857	Sequence 1857, Ap
20	211	15.7	496	14	US-10-106-698-2663	Sequence 2663, Ap
21	97	7.2	549	11	US-09-991-936-1707	Sequence 1707, Ap
22	96.5	7.2	966	11	US-09-974-879-25	Sequence 25, Appl
23	96.5	7.2	966	11	US-09-305-736-25	Sequence 25, Appl
24	96	7.1	3139	12	US-10-252-157-457	Sequence 457, Ap
25	88.5	6.6	1061	14	US-10-161-572-19	Sequence 19, Appl
26	88.5	6.6	1094	14	US-10-161-572-17	Sequence 17, Appl
27	88.5	6.6	1130	14	US-10-161-572-18	Sequence 18, Appl
28	88.5	6.6	1603	10	US-09-966-546-3	Sequence 3, Appl1
29	88.5	6.6	1603	10	US-09-966-545-3	Sequence 3, Appl1
30	88.5	6.6	1603	11	US-09-965-212-3	Sequence 3, Appl1
31	88.5	6.6	1603	15	US-10-189-940-3	Sequence 3, Appl1
32	88.5	6.6	1679	13	US-10-098-841-72	Sequence 72, Appl
33	88.5	6.6	1679	10	US-09-978-295A-522	Sequence 522, App
34	88.5	6.6	1679	10	US-09-978-697-522	Sequence 522, App
35	88.5	6.6	1679	10	US-09-978-192A-522	Sequence 522, App
36	88.5	6.6	1679	10	US-09-999-832A-522	Sequence 522, App
37	88.5	6.6	1679	11	US-09-978-189-522	Sequence 522, App
38	88.5	6.6	1679	11	US-09-978-608A-522	Sequence 522, App
39	88.5	6.6	1679	11	US-09-978-585A-522	Sequence 522, App
40	88.5	6.6	1679	11	US-09-978-191A-522	Sequence 522, App
41	88.5	6.6	1679	11	US-09-978-403A-522	Sequence 522, App
42	88.5	6.6	1679	11	US-09-978-564A-522	Sequence 522, App
43	88.5	6.6	1679	11	US-09-999-833A-522	Sequence 522, App
44	88.5	6.6	1679	11	US-09-981-915A-522	Sequence 522, App
45	88.5	6.6	1679	11	US-09-978-824-522	Sequence 522, App

ALIGNMENTS

RESULT 1
US-09-990-415A-1
; Sequence 1, Application US/09990415A
; Patent No. US20020165182A1
GENERAL INFORMATION:
; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Protein Cluster I
; FILE REFERENCE: 00349
; CURRENT APPLICATION NUMBER: US/09/990,415A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
; LENGTH: 1232
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (450)..(1232)
US-09-990-415A-1

Alignment Scores:
Pred. No.: 3.19e-171
Score: 1345.00
Percent Similarity: 100.00%

Length: 1232
Matches: 261
Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-990-415A-2 (1-261) x US-09-990-415A-1 (1-1232)

```

QY 1 MetGUserLyMeGlyGluLeuProLeuAspIleAsnIleGlnIuProArgrTpaap 20
DB 450 ATGGAAGCAAAATGGGTAATGGCTTTAGACATCAACATCCAGAACCTCGCTGGGAC 509
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgrAsnLeu 40
DB 510 CAAAGACTTCTCGGAGAGAGCCGCGCTTTTACTGCTACTGATCTCGAATAATCTG 569
QY 41 LeuLeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
DB 570 CTGCTGTCGGGGGACAGCTGGAGAGCTTCTCGAACATCTGACAGAACTACAGGGCGGC 629
QY 61 ValValThrProGlyTyleThrGluAspGlnLeuTyrParAlaLysTyrValTyrAspSer 80
DB 630 GTGGTGACCCCGAGGATCACCGAGACAGCTGTGGAGGCCAAGATGTGTATGACTCC 689
QY 81 AlaPheHisProAspThrGlyGlyLysValValLeuIleGlyArgMetSerAlaGlnVal 100
DB 690 GCCTTCATCCGAGACACAGGAGAGAGGTGCTCTGATTTGCCGCTGACGCCAGGTG 749
QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
DB 750 CCCATGAACATACCATCTGCTGATGCTGATCTGATCTGATGATGATGATGATGATGAT 809
QY 121 ValPheTyrPGLTyrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
DB 810 GTTCTTGGCATGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 869
QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
DB 870 GGTGACACTCCCATCTGATGAGGAGAGCTGGGAGAGCCTATGTGATGATGATGATGATGAT 929
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrIleHisLeuProLeuValGly 180
DB 930 GCTGTGGCCAGCGCCCTGGGACTCAAAATCCCTCACCAACACCTGCCCTTGGTGGCG 989
QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
DB 990 AGATTGTGGCCCTTTGGACAGCACTGGCAGCTGCCAATGATCAACATCCCTGATGAGG 1049
QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
DB 1050 CAGAGAGAGCTCAGAGTGGGATCCCGGTGGCTGATGAGGAGGAGGAGGAGGAGGAGGAG 1109
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
DB 1110 TCGGTGACTGACGCAAGCAAGGAAATCTTCCAGGTGGTATTTCAAGAAATCTGATGGCG 1169
QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlyLysLysAspPheLeu 260
DB 1170 ATTCTTGGCATGGCATCCACACTGATGATGATGATGATGATGATGATGATGATGATGAT 1229
QY 261 Lys 261
DB 1230 MAG 1232

```

RESULT 2
 US-09-867-550-1343

Sequence 1343 Application US/09867550

Patent No. US2002008206A1

GENERAL INFORMATION:

APPLICANT: Leach, Martin D.

APPLICANT: Mehrahan, Foad.

APPLICANT: Conley, Pamela

APPLICANT: Law, Debbie

APPLICANT: Topper, James

TITLE OF INVENTION: No. US2002008206A1el Polynucleotides from Atherogenic Cells and

```

; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1343
; LENGTH: 1375
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2)
; OTHER INFORMATION: Wherein any n is one of a or t or c or g
US-09-867-550-1343

Alignment Scores:
Pred. No.: 3,79e-171 Length: 1375
Score: 1345.00 Matches: 261
Best Local Similarity: 100.00% Conservative: 0
Query Match: 100.00% Mismatches: 0
DB: 9 Indels: 0 Gaps: 0

US-09-990-415A-2 (1-261) x US-09-867-550-1343 (1-1375)

QY 1 MetGUserLyMeGlyGluLeuProLeuAspIleAsnIleGlnIuProArgrTpaap 20
DB 91 ATGGAAGCAAAATGGGTAATGGCTTTAGACATCAACATCCAGAACCTCGCTGGGAC 150
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgrAsnLeu 40
DB 151 CAAAGACTTCTCGGAGAGAGCCGCGCTTTTACTGCTACTGATCTCGAATAATCTG 210
QY 41 LeuLeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
DB 211 CTGCTGTCGGGGGACAGCTGGAGAGCTTCTCGAACATCTGACAGAACTACAGGGCGGC 270
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
DB 271 GTGGTGACCCCGAGGATCACCGAGACAGCTGTGGAGGCCAAGATGTGTATGACTCC 330
QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlyLysLysAspPheLeu 260
DB 331 GCCTTCATCCCGAGACACAGGAGAGGAGGTGCTGATGAGGAGGAGGAGGAGGAGGAGGAG 390
QY 261 Lys 261
DB 391 CCCATGAACATACCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 450
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
DB 451 GTGGTGACCCCGAGGATCACCGAGACAGCTGTGGAGGCCAAGATGTGTATGACTCC 510
QY 41 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
DB 511 GGTGACACTCCCATCTGATGAGGAGAGCTGGGAGAGCCTATGTGATGATGATGATGATGAT 570
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrIleHisLeuProLeuValGly 180
DB 571 GCTGTGGCCAGCGCCCTGGGACTCAAAATCCCTCACCAACACCTGCCCTTGGTGGCG 630
QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
DB 631 AGATTGTGGCCCTTTGGACAGCACTGGCAGCTGCCAATGATCAACATCCCTGATGAGG 690
QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
DB 691 CAGAGAGAGCTCAGAGTGGGATCCCGGTGGCTGATGAGGAGGAGGAGGAGGAGGAGGAG 750
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240

```



```
Db 751 TCGGTACTGCACCAAGCAGGGAATCTTCAGGTGGTGTATTCAGAAATCTCATGGCG 810
QY 241 11eProAlaMetAla1leProPLeu1leMetAspThrLeuGluLysAspPheLeu 260
Db 811 ATTCCGCCCATGGCCATCCACACATGATCATGTGACACTGTGAGAGAAAGACTTCTCTG 870
QY 261 Lys 261
Db 871 AAG 873

RESULT 3
US-10-037-270-1016
; Sequence 1016, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhwei
; APPLICANT: Tillinghast, John
; APPLICANT: Dimanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/10/037, 270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: PC_FTL_genes Version 1.0
; SEQ ID NO 1016
; LENGTH: 1729
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (247)..(1371)
US-10-037-270-1016

Alignment Scores:
Pctd. No.: 5,45e-171 Length: 1729
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-990-415a-2 (1-261) x US-10-037-270-1016 (1-1729)
QY 1 MetGluSerLysMetGlyGluLeuProLeuAsp1leAsn1leGlnGluProArgTrpPasp 20
Db 247 ATGGAAAGCAAAATGGGTGATTCCTTAGACATCAACATCCAGAACCTCCCTGGGAC 306
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
Db 307 CAAGGACTTCTCTGGGAGAGCCCGGCACTTTTCACGTACTGATCCTCGAAATCTG 366
QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsn1leValGlnAsnTrpArgAlaGly 60
```

```
Db 367 CTGCTTCGGGGCACAGCTGGAAAGCTTCTCGGAACATCTGCAGAACTACAGGGCCGGC 426
QY 61 ValValThrProGlyIleThrGlyLysAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80
Db 427 GTGGTAGCCCCAGGAGATCACCAGAGACCTGTGGAGGGCCAAAGTATGTGTATGATCC 486
QY 81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db 487 GCCTTCATCCGAGACAGAGGGAGAGAGTGGTCCGTATGCGCGCATGACGCCAGGTG 546
QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysTrpProThrVal 120
Db 547 CCCATGAACATGACCACTCATCTGCTCATCTCTACATTCACAGGAAGACCCCAACCGTG 606
QY 121 ValPheTrpGlnTrpValAlaGlnSerPheAsnAla1leValAsnTyrSerAsnArgSer 140
Db 607 GTGTTCTGGCAGTGGGGAATCACTCTTCAATGCCATTTGTTAACTACTCCACCGGAGT 666
QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
Db 667 GGTGACACTCCCATCACTGTGAGGACAGCTGGGACAGCCATGTGAGTGCACACCTGGA 726
QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrIshHisLeuProPLeuValGly 180
Db 727 GGTGGCCACAGGCTGGGACATCAATCCCTCAACAGCACCTGCCCTTGGTGGC 786
QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
Db 787 AGATTGTGGCTTTGTGAGCAGTGGCAGCTGCCACATGCATCAACATCCCTGATAGG 846
QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgGlyTyr 220
Db 847 CAGAGAGAGCTGACAGTGGGCAATCCCGTGGCTGATGAGCAGAGCTCAGAGGCTTGGCTAC 906
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db 907 TCGGTGACTGCACCAAGCAGGGAATCTTCAGGTGTGATTCAGAAATCTGATGGCG 966
QY 241 11eProAlaMetAla1leProPLeu1leMetAspThrLeuGluLysAspPheLeu 260
Db 967 ATTCCGCCCATGGCCATCCACACATGATCATGTGACACTGTGAGAGAAAGAAAGACTTCTG 1026
QY 261 Lys 261
Db 1027 AAG 1029

RESULT 4
US-10-014-338-1
; Sequence 1, Application US/10014338
; Publication No. US2003092614A1
; GENERAL INFORMATION:
; APPLICANT: Heratb, et al.
; TITLE OF INVENTION: ADP1-41, A NOVEL PROTEIN ISOLATED FROM BRAIN TISSUE HOMOGENAT
; FILE REFERENCE: 9195-077
; CURRENT APPLICATION NUMBER: US/10/014,338
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 10/014,338
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1121)..(1122)
; OTHER INFORMATION: where "n" is any nucleotide
; NAME/KEY: misc.feature
; LOCATION: (1125)..(1126)
; OTHER INFORMATION: where "n" is any nucleotide
```



```

QY 226 LysGlnGlyIlePheGlnValIleSerArgIleCysMetAlaIleProAlaMetAla 245
DB 791 AAACAACCCATCCAGCAAGTTGTCGATCCAGATTCATGCGACGCCCTGGCATGCGC 850
QY 246 IleProIleuIleMetAspThrLeuGluIlyLysAspPheLeuLys 261
DB 851 ATCCCTCCATTCATTATGAACACCTTTGGAAAAGAAACCTTTTGAAG 898

RESULT 6
US-10-198-846-9767
: Sequence 9767, Application US/10198846
: Publication No. US20030099974A1
: GENERAL INFORMATION:
: APPLICANT: Lillie, James
: APPLICANT: Xu, Yongyao
: APPLICANT: Wang, Youzhen
: APPLICANT: Steimann, Kathleen
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
: TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
: TITLE OF INVENTION: THERAPY OF BREAST CANCER
: FILE REFERENCE: MRI-049
: CURRENT APPLICATION NUMBER: US/10/198, 846
: PRIORITY FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/306, 220
: NUMBER OF SEQ ID NOS: 14084
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9767
: LENGTH: 2978
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1, 2, 2974, 2975, 2976, 2977, 2978
: OTHER INFORMATION: n = A,T,C or G
US-10-198-846-9767

Alignment Scores:
Pred. No.: 4,88e-136 Length: 2978
Score: 1089.00 Matches: 201
Percent Similarity: 89.45% Conservative: 28
Best Local Similarity: 78.52% Mismatches: 27
Query Match: 80.97% Indels: 0
Gaps: 0

US-09-990-415a-2 (1-261) x US-10-198-846-9767 (1-2978)
QY 6 GlnGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGlnSerThrPheLeu 25
DB 128 GGAGAACTACCACCAACATTAAACATCAAGAGACCTCGATGGATCAAAAGCATTTCAAT 187
QY 26 GlnArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
DB 188 GGAGAGCCCAATCATTTCTTCACGTGTAACGACCCAGCAAGCAATCTGTTACCAAGAA 247
QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValThrProGly 65
DB 248 CAATTCGAGAGTGGAGAAAAATAGTACATGATTACAGGCAAGAAATTTGTCCTCTGCT 307
QY 66 IleThrGlnAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisProAsp 85
DB 308 CTTCACAAATAATGATTGTTGGAGAGCAAGTACATCTATGATTCAGCTTTTCATCTGAC 367
QY 86 ThrGlnGluLysValIleLeuIleGlyArgMetSerAlaGlnValPrometAsnMetThr 105
DB 368 ACTGCTGAGAAAGATTTTGATAGGAAGATGTCAGGCCAGGTTCCCATCAACATGACC 427
QY 106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValAlaPheTrpGlnTrp 125
DB 428 ATCCAGAGTTGTATGAGGACGTTTACAGGACTGCGCGCTGCTGTGTCGACAGTGG 487
QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
DB 617 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 617

```

```

DB 488 ATTACCAAGCTTCATTAATGCCGCTGCATTAATACCAACAGAACTGGAGCACCCTC 547
QY 146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
DB 548 ACTGTCAATGATGGTGGAGAACGCTTACGTTCTGCACAACTGGTGGCGTGAACAGCT 607
QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProIleuValGlyArgPheValProPhe 185
DB 608 CTAGAGCTCAATGACATTCAGCCAAACATGCTCACCACATGATGAGACGTTGTTCCTCT 667
QY 186 AlaAlaValAlaAlaAsnCysIleAsnIleProIleuMetArgGlnArgGlnLeuGln 205
DB 668 GCTCCGCTAGCTGCTGAATTCATTAATTCATTAATGAGCAAAAGGCACTCAAA 727
QY 206 ValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
DB 728 GTTGCAATTCCTCCGTCAGAGATGAGAAATGGAAACGCTTGGGGAGTGGCGAAGCTCG 787
QY 226 LysGlnGlyIlePheGlnValIleSerArgIleCysMetAlaIleProAlaMetAla 245
DB 788 AAACAACCCATCCAGCAAGTTGTCGATCCAGATTCATGCGACGCCCTGGCATGCGC 847
QY 246 IleProIleuIleMetAspThrLeuGluIlyLysAspPheLeuLys 261
DB 848 ATCCCTCCATTCATTATGAACACCTTTGGAAAAGAAACCTTTTGAAG 895

RESULT 7
US-09-990-415a-3
: Sequence 3, Application US/09990415A
: Patent No. US20020165182A1
: GENERAL INFORMATION:
: APPLICANT: Pharmacia AB
: TITLE OF INVENTION: Protein Cluster I
: FILE REFERENCE: 00349
: CURRENT APPLICATION NUMBER: US/09/990,415A
: PRIORITY FILING DATE: 2001-11-21
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 3
: LENGTH: 1061
: TYPE: DNA
: ORGANISM: human
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (450)..(680)
US-09-990-415a-3

Alignment Scores:
Pred. No.: 3.37e-125 Length: 1061
Score: 1003.50 Matches: 204
Percent Similarity: 78.16% Conservative: 0
Best Local Similarity: 74.61% Mismatches: 57
Query Match: 78.61% Indels: 57
Gaps: 1

US-09-990-415a-2 (1-261) x US-09-990-415a-3 (1-1061)
QY 1 MetGlnSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAsp 20
DB 450 ATGGAAGCAAAATGGGTGAATTCGCTTTACATCAACATCAAGAACTCGCTGGAC 509
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
DB 510 CAAAGTACTTTCCTGGGAGAGCCCGGCACTTTTACATGTTACGTGATCTCGAAACTCG 569
QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
DB 570 CTGCTGTCGGGGACACGCTGGAAGCTTCTGGAACATCGTGCAGAAC----- 617
QY 61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSer 80
DB 617 ----- 617

```

```

QY      81 AlapheHisProAspThrGlyGlyValValLeuIleGlyArgMetSerAlaGlnVal 100
Db      617 -----
QY      101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgIleThrProThrVal 120
Db      618 -----TACAGAGAGACCCCAACCGTG 638
QY      121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db      639 GGTTCCTGGCAGTGGGTGAATCAGTCCATTCATGTCATGTTAACTACTCCAAACCGCAGT 698
QY      141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
Db      699 GGTGACCTCCCATCTGCTGTGAGGACGCTGGGACAGCCCTATGTAGTCCACCACTGGA 758
QY      161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
Db      759 GCTGTGGCCACGGCCCTGGACTCAATCCCTCCACCAAGCACCTCCCCCTTGTGGC 818
QY      181 ArgPheValProPheAlaAlaValAlaAlaAsnGlyIleAsnIleProLeuMetArg 200
Db      819 AGATTGTGCTCTTGCAGCAGTGGCAGCTGCCCACTGCATCAATCAATCCCCCTGTAGAG 878
QY      201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db      879 CAGAGAGAGCTGCAGGTGGCATCCCGGTGCTGTAGGCGAGTCAAGGCTTGGCTAC 938
QY      221 SerValThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAla 240
Db      939 TCGGAGCTGCAGCCCAACAGAGGAAATCTCCAGGTGGTGAATTCAGAAATCGCATGGCG 998
QY      241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
Db      999 ATCTCTGCATGAGCCATCCACCACTGATGACACTGTGAGAAAGAAAGACTTCTCTG 1058
QY      261 Lys 261
Db      1059 AAG 1061

RESULT 8
US-10-014-338-3
; Sequence 3, Application US/10014338
; Publication No. US20030092614A1
; GENERAL INFORMATION:
; APPLICANT: Herath et al.
; TITLE OF INVENTION: ADPT-41, A NOVEL PROTEIN ISOLATED FROM BRAIN TISSUE HOMOGENATE AN
; FILE REFERENCE: 9195-077
; CURRENT APPLICATION NUMBER: US/10/014,338
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 10/014,338
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (949)..(950)
; OTHER INFORMATION: where "n" is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (979)..(980)
; OTHER INFORMATION: where "n" is any nucleotide
US-10-014-338-3

Alignment Scores:
Pred. No.: 6,22e-111 Length: 984
Score: 897.50 Matches: 169
Percent Similarity: 76.26% Conservative: 27

```

```

Best Local Similarity: 65.76% Mismatches: 17
Query Match: 66.73% Indels: 44
DB: 14 Gaps: 1

US-09-990-415a-2 (1-261) x US-10-014-338-3 (1-984)

QY      6 GlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAspGlnSerThrPheLeu 25
Db      17 GGAGAACTACCCACCAACACTTAAATCAAGAAAGAACTCGATGGATCAAAAGCACTTCAATT 76
QY      26 GlyArgAlaArgHisPhePheThrValIleThrAspProArgAsnLeuLeuSerGlyAla 45
Db      77 GGACGAGCCAAATCATTTCTTCACTGTAATGACGCCACCGAGAACATCTGTAAACCAAGAA 136
QY      46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrProGly 65
Db      137 CAACCTCGAGAGCGCCAGAAATAATACATGATTTACAGCAAGAAAGATTTCTCTGCT 196
QY      66 IleThrGluAspGlnLeuTrpArgAlaLysTyrValTyrAspSerAlaPheHisProAsp 85
Db      197 CTTACAGAAATGAATTTGTGGAGACCAAGATACATCTATGATTCACTTTCATCTGC 256
QY      86 ThrGlyGlyLysValValLeuIleGlyArgMetSerAlaGlnValProMetAsnMetThr 105
Db      257 ACTGTTGAGAAAGATATTTGTATGAGAAAGATGTAGCCAGCTTCCATGAACATGACC 316
QY      106 IleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125
Db      317 ATCAGAGGTTGATGATACCGTTTACAGAGATACGCGGCTGTCTGTTCGGCAGTGG 376
QY      126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145
Db      377 ATTAACAGATCTTCAATGACCGCTGCTCAATTAACCAACAGAAAGTGGAGACGACCCCTC 436
QY      146 ThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAla 165
Db      437 ACTGTCATGAGTTGGGAAACAGCTTACCTTCTGTACAAACAGTGGCCGTACACACACT 496
QY      166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
Db      497 CTAGACTCAATGCTGTGACCAAGCATCTCACCACATGATAGAGCGTTTGTCCCTTT 556
QY      186 AlaAlaValAlaAlaAlaAsnGlyIleAsnIleProLeuMetArgGlnArgGluLeuGln 205
Db      557 GCTGCCGTAAGTCTGCTGAATTCATTGAATTCATTAAAGAGGCAA----- 604
QY      206 ValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThrAlaAla 225
Db      604 ----- 604
QY      226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
Db      605 -----AGC 607
QY      246 -IleProProLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
Db      608 CATCCCTTCATTCATTATGAACACTTGGAAAGAAAGACCTTTTGAAG 656

RESULT 9
US-09-990-415a-5
; Sequence 5, Application US/09990415A
; Patent No. US20020165182A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia AB
; TITLE OF INVENTION: Protein Cluster I
; FILE REFERENCE: 00349
; CURRENT APPLICATION NUMBER: US/09/990,415A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1567
; TYPE: DNA

```

```

; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1015)
US-09-990-415A-5

Alignment Scores:
Pred. No.: 3,14e-91 Length: 1567
Score: 754.00 Matches: 150
Percent Similarity: 72.20% Conservative: 37
Best Local Similarity: 57.92% Mismatches: 72
Query Match: 56.06% Indels: 0
DB: 10 Gaps: 0

US-09-990-415A-2 (1-261) x US-09-990-415A-5 (1-1567)

QY 3 SerLyvMetGLyGLuLeuProLeuAspIleasnIleGIngluProArgTrpAspGlnSer 22
DB 41 AGCAAGATGGAGGCTGACCTGTCTGCTTAAACATCGATGCCCGCGTTGGAGCAGCGC 100
QY 23 ThrPheLeuGLyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeu 42
DB 101 ACCCTCGTGGGAGAGTAGAAGCATTCTTAACATCAGAGACCGCCGCGAGCTCTTTGTA 160
QY 43 SerGLyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTrpArgAlaGlyVal 62
DB 161 TCTGAGCGGAGAGCTGAGTGGCCAGAGTGTGTGAGAGAGCAGATGGGGTGTG 220
QY 63 ThrProGLyIleThrGLuAspGlnLeuTrpArgAlaLysTrpValTrpAspSerAlaPhe 82
DB 221 CCCCAGGACCCCAAGTGGAGAGAGCTGTGTATGCCAAGAGCTTATGACTGGCCCTTC 280
QY 83 HisProAspThrGLyGLuLysValValLeuIleGLyArgMetSerAlaGlnValPromet 102
DB 281 CACCCGAGACATGGGAGAGATGATGATCATCGGCGCAGATCTTCCAGCTTCCTGGC 340
QY 103 AsnMetThrIleThrGLyCysMetLeuThrPheTrpArgLysTrpProThrValAlaPhe 122
DB 341 GGCATGATCATCAGCGGCTTCCATGCTCCATCAGAGAGATGCCGCGGTGATCTTC 400
QY 123 TrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTrpSerAsnArgSerGlyasp 142
DB 401 TGGGAGTGGGAGAACCATCTCTCAATGCTTACAGCATACACAGAGATCGGCT 460
QY 143 ThrProIleThrValArgGlnLeuGLyThrAlaTrpValSerAlaThrThrGLyAlaVal 162
DB 461 TCCCCACATCATGCTCAGAGATGGCCCTTCTCTCACTCACAGACACACACATGCTGTG 520
QY 163 AlaThrAlaLeuGLyLeuLysSerLeuThrLysHisLeuProProLeuValGLyArgPhe 182
DB 521 GCCAGGCTGTGGGATGACATGATGTGACAAAGAAAGCCGCGCTGTGGCGGCTGG 580
QY 183 ValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArg 202
DB 581 GTGCCCTTGGCGGTGGCTGTGGCTCACTGTCAATGCCATATCCCATATGCGAGAGG 640
QY 203 GluLeuGlnValGLyIleProValAlaAspGlnAlaGLyGlnArgLeuGLyTrpSerVal 222
DB 641 GAGCTCATAAAGGAGATCGCTGAGAGAGAGATGAATGAATGATGATGATCCCGG 700
QY 223 ThrAlaAlaLysGlnGLyIlePheGlnValValIleSerArgGlyIleCysMetAlaLeuPro 242
DB 701 AGAGCTGGGCGCATRAGGATCATCCCAAGTATTTCTCGATCATCACCATTCTCTCT 760
QY 243 AlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeuLys 261
DB 761 GGGATGATCTTGTGCGCAGTCATCATGAAAGGCTTGAGAAATTCACACTTATGAG 817

```

```

; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungling
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qian A.
; APPLICANT: Ren, Feiyun
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/098,841
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 64
; LENGTH: 2376
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (359)..(1327)
US-10-098-841-64

Alignment Scores:
Pred. No.: 2,1e-90 Length: 2376
Score: 750.00 Matches: 149
Percent Similarity: 72.20% Conservative: 38
Best Local Similarity: 57.53% Mismatches: 72
Query Match: 55.76% Indels: 0
DB: 13 Gaps: 0

US-09-990-415A-2 (1-261) x US-10-098-841-64 (1-2376)

QY 3 SerLyvMetGLyGLuLeuProLeuAspIleasnIleGIngluProArgTrpAspGlnSer 22
DB 353 AGCAAGATGGAGGCTGACCTGTCTGCTTAAACATCGATGCCCGCGTTGGAGCAGCGC 412
QY 23 ThrPheLeuGLyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeu 42
DB 413 ACCCTCGTGGGAGAGTAGAAGCATTCTTAACATCAGAGACCGCCGCGAGCTCTTTGTA 472
QY 43 SerGLyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTrpArgAlaGlyVal 62
DB 473 TCTGAGCGGAGAGCTGAGTGGCCAAAGTGTGTGAGAGAGAGATGGGGTGTG 532
QY 63 ThrProGLyIleThrGLuAspGlnLeuTrpArgAlaLysTrpValTrpAspSerAlaPhe 82
DB 533 CCCCAGGACCCCAAGTGGAGAGCTGTGTATGCCAAGAGCTGTATGACTGGCCCTTC 592
QY 83 HisProAspThrGLyGLuLysValValLeuIleGLyArgMetSerAlaGlnValPromet 102
DB 593 CACCCGAGACATGGGAGAGATGATGATCATCGGCGCATGTCTTCCAGCTTCTGGC 652
QY 103 AsnMetThrIleThrGLyCysMetLeuThrPheTrpArgLysTrpProThrValAlaPhe 122
DB 653 GGCATGATCATCAGCGGCTTCACTGCTCCAGTTCACAGAGAGATGCCGCGGTGATCTTC 712
QY 123 TrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTrpSerAsnArgSerGlyasp 142

```

```

RESULT 10
US-10-098-841-64
; Sequence 64, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:

```

```

Db      713 TGGCAGTGGTGAACACATCCTTCAATGCTTACAGCAACACAGGAAATCGGCT 772
Qy      143 ThrProIleThrValArgGlnLeuGlyThrAlaTyValSerAlaThrThrGlyAlaVal 162
Db      773 TCCCCACATCAGTCAGGACAGATGCGCTTCCCTACTCCACAGCACAACCACTGCTGTG 832
Qy      163 AlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPhe 182
Db      833 GCCACGCTGTGGGCATCAACATGTTGACAAAGAAAGCCGCCCTGTGGTGGCGCTGTG 892
Qy      183 ValProPheAlaAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 202
Db      893 GTGCCCTTGGCGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTG 952
Qy      203 GlnLeuGlnValGlyIleProValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 222
Db      953 GAGCTCATMAAGGAAATCGCTGTGAAGGACAGAAATGAAATGATGATGATGATGATG 1012
Qy      223 ThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAlaAlaPro 242
Db      1013 AGAGCTGGCCCATAGGCATCACTCACTAGTATTCTCGATCACCATGTCACCTCTCT 1072
Qy      243 AlaMetAlaAlaProProLeuIleMetAspThrLeuGlnLysLysAspPheLeuLys 261
Db      1073 GGATGATCTTGTGCTGCCATCATGTGAAGGCTTGAAGAAATTGCACCTTATGACG 1129

RESULT 11
US-09-822-849A-490
; Sequence 490, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822, 849A
; PRIORITY FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 490
; LENGTH: 2458
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-490

Alignment Scores:
Pred. No.: 2,22e-90 Length: 2458
Score: 750.00 Matches: 149
Percent Similarity: 72.20% Conservative: 38
Best Local Similarity: 57.53% Mismatches: 72
Query Match: 55.76% Indels: 0
DB: 9 Gaps: 0

US-09-990-415A-2 (1-261) x US-09-822-849A-490 (1-2458)

Qy      3 SerLysMetGlyGlnLeuProLeuAspIleAsnIleGlnLysProArgTyrPaspGlnSer 22
Db      75 AGCAAGATGAGAGGCTGACCTGTCTGCTTAACATCGATGCCCGCTGTGGAGACACGCG 134
Qy      23 ThrPheLeuGlyArgAlaArgHisPhePheThrValLhrAspProArgAsnLeuLeu 42
Db      135 ACCTTCTCTGGGAGAGTGAAGCACTTCTTAACATCAAGGACCCCGCCTGCTCTTTGTA 194

```

```

Qy      43 SerGlyAlaGlnLeuGlnLysSerArgAsnIleValGlnAsnTyrArgAlaGlyAlaVal 62
Db      195 TCTGAGCGGAGAGCTGGATGGCCCAAGTGAATGCTGAGAGAGACGATGGGGTGTG 254
Qy      63 ThrProGlyIleThrGlnAspGlnLeuThrPargAlaLysTyrValTyrAspSerAlaPhe 82
Db      255 CCCCCAGCACCCAAAGTAGAGAGAGCTGCTGATGCGCAAAAAGCTGTATGACTGCGCTTC 314
Qy      83 HisProAspThrGlyGlnLysValAlaValLeuIleGlyArgMetSerAlaGlnValProMet 102
Db      315 CACCCGACACTGGGGAGAGATGAAATGCAATCGGGCCAGATGCTTTCACGCTTCTGCG 374
Qy      103 AsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValAlaPhe 122
Db      375 GGCATGATCAACACGCGCTTCCATGCTTCCAGAGAGATCCCGCGGTGATCTTC 434
Qy      123 TrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAspArgSerGlyasp 142
Db      435 TGGCAGTGGTGMACAGTCTTCAATGCTTACACTACACCAACAGGAAATCGGCT 494
Qy      143 ThrProIleThrValArgGlnLeuGlyThrAlaTyValSerAlaThrThrGlyAlaVal 162
Db      495 TCCCCACATCAGTCAGGACAGATGGCCCTTCTCTCACTTCAACAGCACAACCACTGCTG 554
Qy      163 AlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPhe 182
Db      555 GCCACGCTGTGGGCATCAACATGTTGACAAAGAAAGCCGCCCTGTGGTGGCGCTGTG 614
Qy      183 ValProPheAlaAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 202
Db      615 GTGCCCTTGGCGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTG 674
Qy      203 GlnLeuGlnValGlyIleProValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 222
Db      675 GAGCTCATMAAGGAAATCGCTGTGAAGGACAGAAATGAAATGATGATGATGATGATG 734
Qy      223 ThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAlaAlaPro 242
Db      735 AGAGCTGGCCCATAGGCATCACTCACTAGTATTCTCGATCACCATGTCACCTCTCT 794
Qy      243 AlaMetAlaAlaProProLeuIleMetAspThrLeuGlnLysLysAspPheLeuLys 261
Db      795 GGATGATCTTGTGCTGCCATCATGTGAAGGCTTGAAGAAATTGCACCTTATGACG 851

RESULT 12
US-10-206-901B-41
; Sequence 41, Application US/10206901B
; Publication No. US20030100540A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, ZHONGHUA
; APPLICANT: DUBOIS, RAYMOND
; TITLE OF INVENTION: IDENTIFICATION OF NSAID-REGULATED GENES
; FILE REFERENCE: VBLT:01205
; CURRENT APPLICATION NUMBER: US/10/206, 901B
; PRIORITY FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/308,370
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 670
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-206-901B-41

Alignment Scores:
Pred. No.: 1,34e-85 Length: 670
Score: 708.00 Matches: 135
Percent Similarity: 88.62% Conservative: 13
Best Local Similarity: 80.84% Mismatches: 19
Query Match: 52.64% Indels: 0
DB: 14 Gaps: 0

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: October 16, 2003, 15:58:12 ; Search time 2075 Seconds

(without alignments)
3057.093 Million cell updates/sec

Title: US-09-990-415A-2
Perfect score: 1345
Sequence: 1 MESKMGELPDLINIOEPKWD.....PAMAIPLINDLEKDFLK 261

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+.p2n.model -DEV=xlp
-O=/sgn2.1/USPRO.spool.p/US09990415/runat.16102003.073807.9449/app.query.fasta.1.455
-DB=ESF -QFWT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTWT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09990415 -SCGN_1_1_2814 -runat.16102003.073807.9449 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESUBSTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vil:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1341	99.7	1201	13	BX376574
2	1333	99.1	869	13	BQ223306
3	1325	98.5	975	13	BQ933752
4	1312	97.5	998	12	BM473966
5	1305	97.0	1201	13	BX34089
6	1291	96.0	1076	12	BM925231
7	1290	95.9	995	13	BX355800
8	1283	95.4	898	14	CB203758
9	1283	95.4	2682	11	AK034514
10	1216	90.4	963	13	BQ959114
11	1212	90.1	934	13	BQ506082
12	1196	88.9	1040	12	BM548674
13	1189	88.4	961	12	BG752086
14	1180	87.7	862	13	BQ425697
15	1166	86.7	731	14	CA317398
16	1166	86.7	811	12	B1818854
17	1166	86.7	870	12	BM047107
18	1162	86.4	764	14	CA317392
19	1153	85.7	715	14	CA316261
20	1148.5	85.4	868	10	BE780804
21	1121	83.3	1008	10	BE569711
22	1114	82.8	755	12	B1517760
23	1105	82.2	727	14	CD354095
24	1104.5	82.1	1071	10	BE569784
25	1090.5	81.1	758	10	BF346247
26	1089	81.0	1201	13	BF363487
27	1084.5	80.6	746	12	B1752914
28	1077.5	80.1	2326	11	AK089985
29	1076	80.0	1080	9	AL583707
30	1073	79.8	869	14	CA489784
31	1072	79.7	830	13	BQ408998
32	1071.5	79.7	666	14	CA318631
33	1069	79.5	864	13	BQ435963
34	1059	78.7	1228	11	AK012650
35	1053	78.3	1016	14	BY711785
36	1048.5	78.0	956	12	BM928783
37	1048	77.9	955	13	BQ254199
38	1041	77.4	791	13	BQ355128
39	1039	77.2	786	12	B1152915
40	1036	77.0	1004	12	BM471040
41	1033	76.8	676	12	BM944910
42	1032.5	76.8	782	14	CB234292
43	1032	76.7	655	10	BB626574
44	1032	76.7	793	13	BQ355445
45	1017	75.6	944	12	B1903936

ALIGNMENTS

RESULT 1
LOCUS BX376574 1201 bp mRNA linear EST 08-MAY-2003
DEFINITION BX376574 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
ACCESSION BX376574
VERSION BX376574.1 GI:30448513
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1201)

AUTHORS Ll.W.B., Gruber,C., Jesse,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT

Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqef@genoscope.cns.fr Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5549.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgl-1b1/cluster.cgi?seq=CS0DD001DC030P1ecluster=5549.f. Contact :

Feng Liang Email : fliang@life.techn.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DD001DC030P1.

FEATURES

Source : Location/Qualifiers

1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DD001P06"
 /tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 268 a 306 c 347 g 250 t 30 others
 ORIGIN

Alignment Scores:

Pred. No.: 1.41e-147 Length: 1201
 Score: 1341.00 Matches: 260
 Percent Similarity: 99.62% Conservative: 0
 Best Local Similarity: 99.62% Mismatches: 1
 Query Match: 99.70% Indels: 0
 DB: 13 Gaps: 0

US-09-990-415a-2 (1-261) x BX376574 (1-1201)

1 MetGuseRlysmetGlyglulenuProleuspiIleasnIlegIngluProArqTrrasp 20
 113 ATGGAAACCAAAATGGTGAATGGCTTTAGACATCAACATCCAGAACTCTCGCGGAGC 172
 21 GlnserThrphelLeuGlyArgAlaArgHisphethrValThrAspProArqasnleu 40
 173 CAAGTACTTCTCTGGGAGAGCCCGGACATTTTCACTGTTACTGATCTCGAATCTG 232
 41 LeuLeuSerGlyAlaGlnLeuGlnAlaSerArgasnIleValGlnAsnTyrArgAlaGly 60
 233 CTGCGTCCGGGGCGACACTGGAAGCTTCTCGGAACATCGTGCAGAACTACAGGGCGGCG 292
 61 ValValThrProGlyIleThrGlnAspGlnLeuThrPaArgAlaLysTyrValTyrAspSer 80
 293 GTGGGACCCCAAGGATACCGAGGACCAAGCTGTGGAGGCCAAGTATGTATGATACCC 352
 81 AlaPhenIleProAspThrGlyGlyIuLysValValLeuIleGlyArgMetSerAlaGlnVal 100
 353 GCCTTCCATCCGGGACAGAGGGGAGAAGTGTCTCGATGGCCCGCATGTYAGCCAGGTG 412
 101 PromethasmetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 413 CCCATGAACATGACCATCATCTGCTGCTCATCTACAGTCTACAGGAGAGAGCCCAACGCTG 472
 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 473 GTGTTTCGGAGTGGGTGAATCATGCTTTCATATGCACTGTATTACTACTCCACCGCAGT 532
 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
 533 GGTGACACTCCCATCACTGTGAGGAGAGCTGGGAGAGCTGTGAGAGTCCACCACTGGA 592
 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180

|||||
 593 GCTGTGCCACAGCCCTGGAGCTCAATCCCTACCAAGCAGCTCCCTGGTGGCGC 652
 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnIleAsnIleProLeuMetArg 200
 653 AGATTGTGGCCCTTTGCAGACACTGGCAGCTGCCAAGTGCATCAACATCCCTGATGAG 712
 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuTyr 220
 713 CAGAGAGAGCTGCAGCTGGCATCCCGGTGCTGATGAGGACGTCAGAGCTTGCTAC 772
 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValIleSerArgIleCysMetAla 240
 773 TCGGTGACTGCAGCCAGAGGGAATCTTCCAGGTGTGATTCAGAAATCTGCATGGCG 832
 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysAspPheLeu 260
 833 ATTCCTGCAGTGGCATCCACCACTGATCATGACGACACTCTGAGAAAGACCTTCTG 892
 261 Lys 261
 893 AAG 895

RESULT 2
 BQ223306
 LOCUS BQ223306
 DEFINITION BQ223306 869 bp mRNA linear EST 02-MAY-2002
 AGENCOURT.7518440 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:606572
 5', mRNA sequence.
 ACCESSION BQ223306.1 GI:20404706
 VERSION BQ223306.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 869)
 NIH-MGC http://mhc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL13344 row: k column: 21
 High quality sequence stop: 688.
 Location/Qualifiers

FEATURES

Source : Location/Qualifiers

1..869
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:606572"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_92"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site.1: NotI;
 Site.2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 193 a 244 c 255 g 175 t 2 others
 ORIGIN

Alignment Scores:

Pred. No.: 7.79e-147 Length: 869
 Score: 1333.00 Matches: 259
 Percent Similarity: 99.62% Conservative: 1
 Best Local Similarity: 99.23% Mismatches: 1
 Query Match: 99.11% Indels: 0


```

DB: 13 Gaps: 0
US-09-990-415a-2 (1-261) x B0223306 (1-869)

OY 1 MetGuserLyMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpasp 20
   |||||||
DB 37 ATGGAACCAAAATGGGTGAATGCTTTAGACATCAACATCCAGAACCTCGCGGAGC 96
OY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
   |||||||
DB 97 CAAAGTACTTCTCTGGCAGAGCCCGGACATTTTCACTGTTACTGATCCTCGAATCTG 156
OY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
   |||||||
DB 157 CTGCTGTCGGGGGACACAGCTGGAACTTCTCGGAACATCGTCCAAACTACAGGCGCCGCG 216
OY 61 ValValThrProGlyIleThrGluAspGlnLeuThrParGAlaLysTyrValTyrAspSer 80
   |||||||
DB 217 GTGGTGACCCCAAGGAGATCACCAGAGACCACTGTGGAGGCCCAAGTATGTATGATCTCC 276
OY 81 AlaPheHisProAspThrGlyGlyLysValValLeuIleGlyArgMetSerAlaGlnVal 100
   |||||||
DB 277 GCCTTCATCCGGACACAGGGGAGAGAGGTGGTCCGTGATGGCCGATCCAGCCAGGTG 336
OY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
   |||||||
DB 337 CCCATGAACATGACCATCATCTGGCTGCATCTCATCATCTACAGGAGAACCCCAACCGTG 396
OY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
   |||||||
DB 397 GTGTCTGGCAGGTGGGTAAATCAGCTTCAATGCAATGTTAACTACTCCAAACCGCAGT 456
OY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
   |||||||
DB 457 GGTGACACTCCCATCATCTGTGAGGAGAGCTGGGGAGAGCCTATGTAGTCCACCACTGGA 516
OY 161 AlaValAlaThrAlaLeuGlyLysSerLeuThrLysHisLeuProProLeuValGly 180
   |||||||
DB 517 GCTGTGGCCAGCGGCGCTGGAGCTCAAAATCCCTCACCAAGCACCTCCCGCTGGTGGCG 576
OY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnGlySerIleAsnIleProLeuMetArg 200
   |||||||
DB 577 AGATTGTGCCCTTGTGACAGCACTGGCAGCTGCCAATCATCAATCCCCCTGATGAGG 636
OY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
   |||||||
DB 637 CAGAGAGAGCTGACAGTGGCATCCCGGTGGCTGTAGGAGGAGTGCAGAAAGCTTGGCTAC 696
OY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
   |||||||
DB 697 TCGGTGACTGCAGCCCAAGCAGGGAATCTTCCAGGTGGTGAATTTCAAGATCTGCTGGCG 756
OY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeu 260
   |||||||
DB 757 ATTCTGCGCATGGCCATCCCAAGCACTGATCATGACACTCTGGAGAAAGAACTTCTTG 816
OY 261 Lys 261
   |||
DB 817 AAG 819

RESULT 3
B0933752 975 bp mRNA linear EST 21-AUG-2002
LOCUS B0933752 8837545 NTI_MGC_18 Homo sapiens cDNA clone IMAGE:6424845
DEFINITION 5', mRNA sequence.
ACCESSION B0933752
VERSION B0933752.1 GI:22349135
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 975)

```

```

AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
        Email: cgeabs-remail.nih.gov
        Tissue Procurement: DCTD/DTF/Gazdar
        cDNA Library Preparation: Rubin Laboratory
        cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
        DNA Sequencing by: Agencourt Bioscience Corporation
        Clone distribution: MGC clone distribution information can be
        found through the I.M.A.G.E. Consortium/LNLN at:
        http://image.llnl.gov
        Plate: L1CM2605 row: k column: 22
        High quality sequence stop: 646.
        Location/Qualifiers
            1..975
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6424845"
            /tissue_type="large cell carcinoma"
            /lab_host="DH10B (phage-resistant)"
            /clone_id="NIH_MGC_18"
            /note="Organ: Lung; Vector: pOTB7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCAAGAG(G). Library constructed by Ling Hong in the
            laboratory of Gerald M. Rubin (University of California,
            Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
            Superscript II RT (Life Technologies). Note: this is a
            NIH_MGC Library."
            NIH_MGC Library."

BASE COUNT 225 a 277 c 267 g 199 t 7 others
ORIGIN
Alignment Scores:
Pred. No.: 8,11e-146 Length: 975
Score: 1325.00 Matches: 256
Percent Similarity: 99.61% Conservative: 2
Best Local Similarity: 98.84% Mismatches: 1
Query Match: 98.51% Indels: 0
DB: 13 Gaps: 0

US-09-990-415a-2 (1-261) x B0933752 (1-975)

OY 1 MetGuserLyMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpasp 20
   |||||||
DB 30 ATGGAACCAAAATGGGTGAATGCTTTAGACATCAACATCCAGAACCTCGCGGAGC 89
OY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
   |||||||
DB 90 CAAAGTACTTCTCTGGCAGAGCCCGGACATTTTCACTGTTACTGATCCTCGAATCTG 149
OY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
   |||||||
DB 150 CTGCTGTCGGGGGACACAGCTGGAACTTCTCGGAACATCGTGCAGAACTACAGGCGCGC 209
OY 61 ValValThrProGlyIleThrGluAspGlnLeuThrParGAlaLysTyrValTyrAspSer 80
   |||||||
DB 210 GTGGTGACCCCAAGGAGATCACCAGAGACCACTGTGGAGGCCCAAGTATGTATGATCTCC 269
OY 81 AlaPheHisProAspThrGlyGlyLysValValLeuIleGlyArgMetSerAlaGlnVal 100
   |||||||
DB 270 GCCTTCATCCGGACACAGGGGAGAGAGGTGGTCCGTGATGGCCGATCCAGCCAGGTG 329
OY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
   |||||||
DB 330 CCCATGAACATGACCATCATCTGGCTGCATCTCATCATCTACAGGAGAACCCCAACCGTG 389
OY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
   |||||||
DB 390 GTGTCTGGCAGGTGGGTAAATCAGCTTCAATGCAATGTTAACTACTCCAAACCGCAGT 449
OY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160

```

|||||
Db 450 GGTGACACTCCCATCAGTGTGAGGACGCTGGGAGACGCTTGTGATGAGTCCACCACTGGA 509
Oy 161 AAlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGly 180
Db 510 GCTGGGCGACGCGCTGGGACTCAATCCCTCACAAGACAGCTCCCTGGTGGTGGC 569
Oy 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnGlyHisAsnIleProLeuMetArg 200
Db 570 AGATTGTGCTCCCTTGGACGAGTGGCAGCTCCCACTGATCAACATCCCTCGATGAGG 629
Oy 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db 630 CAGAGAGAGCTGACAGGTGGGATCCCGGTGGCTGATGAGCAGGTCAGAGAACTGGCTAC 689
Oy 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db 690 TCGGAGCTGACGACCAAGCAGGGAATCTCCAGGGTGGATTTCAAGATCTGCATGGCG 749
Oy 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPhe 259
Db 750 ATTCCTCCGATGCGCATCCACCACTGATCATGTGACACTGTGAGAAAGAACTTC 806

RESULT 4 998 bp mRNA linear EST 05-FEB-2002
LOCUS BM473966
DEFINITION AGENCOURT.6489827 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5521420
ACCESSION BM473966
VERSION BM473966.1 GI:18523008
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
Cloned by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
Plate: L1AM12187 row: a column: 05
High quality sequence stop: 652.
Location/Qualifiers
1..998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5521420"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: PCMV-SPORE6; Site: 1: NotI;
Site-2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT 228 a 292 c 279 g 198 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 2.89e-144 Length: 998
Score: 1312.00 Matches: 259
Percent Similarity: 98.86% Conservative: 1
Best Local Similarity: 98.48% Mismatches: 1
Query Match: 97.55% Indels: 2
Gaps: 0

US-09-990-415a-2 (1-261) x BM473966 (1-998)
Oy 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
Db 43 ATGGAAAGCAAAATGGGTGAATGGCTTTAGACATCAACATCCAGGAACCTCGGTGGAGC 102
Oy 21 GlnSerThrPheLeuGlyValArgHisPhePheThrValIleAspProArgAsnLeu 40
Db 103 CAAAGTACTTTCCTGGGCGAGAGCCCGGACCTTTTCACTGTACTGATGACCTCGAAATCTG 162
Oy 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValIleAsnTyrArgAlaGly 60
Db 163 CTGCTGTCGGGGGACAGACTGGAAGCTTCTGGAACATCTGCGAAGACTACAGGCGCGC 222
Oy 61 ValValThrProGlyIleThrGluAspGlnLeuThrParGalaLysTyrValTyrAspSer 80
Db 223 GTGGTGAACCCAGGAGATCACCGAGACAGCTGTGGAGCGCAAGTATGTATGACTCC 282
Oy 81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db 283 GCTTCCATCCGACACAGGGGAGAGAGGTGGTCTGATTTGGCCCATGTCCAGCCAGGTG 342
Oy 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db 343 CCCATGAACATGACATCATCTGCTGCTGCTCATCTTCACAGGAACACCCCAACCTG 402
Oy 121 ValPheThrGlnIleThrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db 403 GTGTTCTGGCAGTGGGTGAATCACTCTTCAATGCCATTTGTAATCTACTCCACCGCAGT 462
Oy 141 GlyAspThrProIleThrValArgGlnLeuGlyThrValArgLysValSerAlaThrThrGly 160
Db 463 GGTGACACTCCCATCAGCTGTAGGACAGCTGGGAGACGCTTGTGAGTCCACCACTGGA 522
Oy 161 AAlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGly 180
Db 523 GCTGGGCGACGCGCTGGGACTCAATCCCTCACAAGACAGCTGCGCCCTGGTGGC 582
Oy 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnGlyHisAsnIleProLeuMetArg 200
Db 583 AGATTGTGCGCTTGGACGAGTGGCAGCTCCCACTGCAACATCCCTCATATGAGG 642
Oy 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db 643 CAGAGAGAGCTGCGAGGTGGGATCCCGGTGGCTGATGAGGAGGTCAGAACTGGCTGAC 702
Oy 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db 703 TCGGTGACTGACGACCAAGCAGGGAATCTTCCAGGTGGGTGATTTCAAGATCTGCATGG 762
Oy 240 IleIleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPhe 260
Db 763 CGATTCTCGCATGNCATCCACCACTGATCATGACACTGTGAGAGAAAGAACTTCC 822
Oy 260 eulys 261
Db 823 TGAAG 827

RESULT 5 1201 bp mRNA linear EST 02-MAY-2003
LOCUS BX334089
DEFINITION BX334089 Homo sapiens NEUROBLASTOMA COR 50-NORMALIZED Homo sapiens
CDNA clone CS0DD006TYP08 5'-PRIME, mRNA sequence.
ACCESSION BX334089
VERSION BX334089.1 GI:30341296
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization

JOURNAL COMMENT Unpublished
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5549.f for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DD006DH04QP&cluster=5549.f. Contact :
 Peng Liang Email : filiang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DD006DH04QP1.

FEATURES
 source
 1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DD006YPO8"
 /issue_type="NEUROBLASTOMA COT 50-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 266 a 316 c 345 g 245 t 29 others

ORIGIN

Alignment Scores:
 Pred. No.: 2.54e-143 Length: 1201
 Score: 1305.00 Matches: 258
 Percent Similarity: 98.85% Conservative: 0
 Best Local Similarity: 98.85% Mismatches: 3
 Query Match: 97.03% Indels: 1
 DB: 13 Gaps: 0

US-09-990-415a-2 (1-261) x BX334089 (1-1201)

QY 1 MetGluSerLysMetGlyLeuProLeuAspPheAsnIleGlnGluProArgTyrPasp 20
 DB 88 ATGGAAGCAAAATGGGTGARTTGCTTGTAGCATCAACATCCAGGAACCTGCTGGAG 147
 QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
 DB 148 CAAGAGTACTTCTGGGACAGAGC-CGGCACTTTTCACTGTACTGATCTCTCGARATCTG 206
 QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 DB 207 CTGCTGTCCGGGGACAGCTGGAAGCTTCTCGAACAATCGTCAGAACTACAGGGCCGGC 286
 QY 61 ValValThrProGlyIleThrGluAspGlnLeuTyrPargAlaAlaLysTyrValTyrAspSer 80
 DB 267 GTGGTACACCCAGGAGATCACCGGAGACCACTGTGGAGGCGCAAGTATGTATGACTCC 326
 QY 81 AlaPheHisProAspThrGlyGlyLysValValLeuIleGlyArgMetSerAlaGlnVal 100
 DB 327 GCGTTCATCCGGACACAGGGGAGAGGTGCTCTATGCGCCGATGTCCAGCCAGGTG 386
 QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 DB 387 CCCATACATGACATCACTGCTGCTCATCTCACAATCCAGGAAGACCCCAACCGTG 446
 QY 121 ValPheTrpGlnTyrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 DB 447 GTGTTTGGAGAGGGGTGATCAGTCTTCATATGCAATGCTTAATCTCAACCCAGT 506
 QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrTrpGly 160
 DB 507 GGTGACACTCCATCTGTGTGAGGACAGGGGACAGCCATGTGTGATGTCACACTGTGA 566
 QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
 DB 567 GCTGTGGCCACGGCCCTGGGAGCTCAATATCCCTCACCAGACACTGCCCTTGTCGCGC 626

QY 181 ArgPheValProPheAlaAlaValAlaAlaAsnCysIleAsnIleProLeuMetArg 200
 DB 627 AGATTTCGCCCTTTCAGACAGTGGACGTCCCACTGTCATCAACATCCCCCTTATAGC 686
 QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
 DB 687 CAGAGAGAGCTGACAGTGGGCAATCCCGTGGCTGATGAGCGAGTCAAGGCTTGCGTAC 746
 QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
 DB 747 TCGGTGACTCAGCAGCAAGAGGAAATCTTCAGGTGTGATTCAGAAATCTGCATGCGC 806
 QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeu 260
 DB 807 ATTCTTCCTCATGGCATCCCACTCCATCATCATGACACTCTGGAGAAAGAACTCTCTG 866
 QY 261 Lys 261
 DB 867 AAG 869

RESULT 6
 BM925231
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1076)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLAM12812 row: p column: 13
 High quality sequence start: 4
 High quality sequence stop: 664.

FEATURES
 source
 1. 1076
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5762556"
 /lab_host="DH10B"
 /clone_lib="NIH-MGC_122"
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
 site.1: NotI; site.2: EcoRV (destroyed); RNA source
 anonymous pool of 24 week female spleens. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 026. Note:
 this is a NIH-MGC library."

BASE COUNT 230 a 316 c 314 g 216 t

ORIGIN

Alignment Scores:
 Pred. No.: 9.77e-142 Length: 1076
 Score: 1291.00 Matches: 255

Db 571 GCWGTGGCCACGGCCCTGGGACCTCAATCCCTCACCAGACCTGCCCCCTTGCTGGC 630
Qy 181 ArpPheValProPheAlaAlaValAlaAlaAlaAsnGlyLeuLeuMetArg 200
Db 631 AGATTTGTGCTTGTGGACAGTGGCAGCTGCCAAGCATCAACATCCCTGTATGAGG 690
Qy 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluLeuGlnArgLeuGlyTyr 220
Db 691 CAAAGAGCTGACGGGGGCGATCCGCTGTATGAGGAGGTCAGAGCTTGGCTTAC 750
Qy 221 SerValThrAlaAlaAlaGlyGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db 751 TCGGTGCTGACGACCAAGGAGGAGATCTTCAGGTGCTGATTCACAATCTGATGCGC 810
Qy 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnIleGlyAspPheLeu 260
Db 811 ATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 870
Qy 261 Lys 261
Db 871 AAG 873
RESULT 8
LOCUS CB203758 898 bp mRNA linear EST 05-FEB-2003
DEFINITION AGENCOURT_11288629 NIH_MGC_135 Mus musculus cDNA clone
IMAGE:30142456 5', mRNA sequence.
ACCESSION CB203758
VERSION CB203758.1 GI:28240427
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL 1 (bases 1 to 898)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: sgabbs@emall.nih.gov
Tissue Procurement: Dr. David Rowe
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: NDAM0051 row: 1 column: 17
High quality sequence stop: 676.
Location/Qualifiers
1. 898
FEATURES
source
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30142456"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_135"
/note="Vector: PCWVSport6.1; Site_1: EcoRV; Site_2: NotI;
Normalized full-length enriched library from pooled mouse
embryonic limb, maxilla and mandible, day 12.5, 13.5, 14.5
, and 15.5 (size selected for the 0.5-1 kb fragments)
Cloned directionally, priming method: oligo-dT. cDNA
enrichment: >1k bp. Average insert size 1.0k bp.
Normalization (cot value): 7.5 kb. Priming sequence:
5'GACTGATCTGATGATGCGGAGCGCGCC(7)3' Tissue contributed by
, David Rowe. Library constructed by Resgen, Invitrogen
Corp."
BASE COUNT 190 a 255 c 253 g 199 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 6.64e-141 Length: 898
Score: 1283.00 Matches: 248

Percent Similarity: 98.08% Conservative: 7
Best Local Similarity: 95.38% Mismatches: 5
Query Match: 95.39% Indels: 0
DB: 14 Gaps: 0
US-09-990-415a-2 (1-261) x CB203758 (1-898)
Qy 2 GlnSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnIleProArgTyrPaspGln 21
Db 13 GAAAGCAAGATGGGTGACCTGCCCTTAAATCAACATCCAGCAAGACCTGGTGACCAA 72
Qy 22 SerThrPheLeuGlnArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeu 41
Db 73 AGCACATTTCTAGGACAGAGCCGATTTCTTCACTGACTGATCTCCAAATCTGCTG 132
Qy 42 LeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyVal 61
Db 133 CTGTGCGGGGAAACAGCTGGAAGCTTCCCGGAACATGTCAGAGATTACAGGCGTGTG 192
Qy 62 ValThrProGlyIleThrGluAspGlnLeuThrPargAlaLysTyrValTyrAspSerAla 81
Db 193 GCAACCCCGGGTCTCACTGAGAGACCACTATGCGGACCAAAATCGTGTATGACTCAGCA 252
Qy 82 PheHisProAspThrGlyGluValValLeuIleGlyArgMetSerAlaGlnValPro 101
Db 253 TTCCATCCGAGACAGGGGGAGAGGTGCTGTTGATGGCCGATGTACGCCAGGTGCC 312
Qy 102 MetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValVal 121
Db 313 ATGAAACATGACCAATTACTGCTGATGCTCACTCACTCAACAGAAAGACTCCGACTGTGTG 372
Qy 122 PheThrProlIleThrValAlaGlnIleSerPheAsnAlaIleValAsnTyrSerAsnArgSerGly 141
Db 373 TTCTGGAGAGGTGCAATCACTCTTCAATGCTATTTGATTAATCTTAATCGCAGCGGC 432
Qy 142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAla 161
Db 433 GATGCTCCATCTCACTGTGCAAGCACTTGGGACAGCCATATGTAGATGCCACACAGGGGCT 492
Qy 162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysIleLeuProProLeuValGlyArg 181
Db 493 GTGGCTACTGCTCGGGAGCTCAAGTCTCTCAACCAAGCACTGCCCGCTAGTCGGTCA 552
Qy 182 PheValProPheAlaAlaValAlaAlaAlaAsnGlyLeuLeuMetArgGln 201
Db 553 TTCTGCTGCTTGTGAGCTGTGGCGCTGCAACATGCTCAACATCCCTGTATGAGGACG 612
Qy 202 ArgGluLeuGlnValGlyIleProValAlaAspGluLeuGlnArgLeuGlyTyrSer 221
Db 613 AGGAGAGCTGACAGGTGGGATCCAGTACATGATGAGCTGTGTCAGAGGCTTGGCCACTCG 672
Qy 222 ValThrAlaAlaAlaGlyGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIle 241
Db 673 GTGACTGCTGCCAAGAGGAAATCTTCAGGTGCTGATTCAGAAATCGGAATGCGCATTT 732
Qy 242 ProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnIleGlyAspPheLeuLys 261
Db 733 CCGGCAATGGCCATTCCTCCCGGTGATCATGAAACATCTGGAGAGAAAGAACTTCTGAAG 792
RESULT 9 AK034514 2682 bp mRNA linear HTC 05-DEC-2002
LOCUS AK034514
DEFINITION Mus musculus adult male diencephalon cDNA, RIKEN full-length
enriched library, clone:9330209L13 product:stideroflexin 3, full
insert sequence.
ACCESSION AK034514
VERSION AK034514.1 GI:26330000
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1

AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Mech. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
REFERENCE	10349636
AUTHORS	2
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)
PUBMED	20493374
REFERENCE	11042159
AUTHORS	3
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
JOURNAL	RIKEN integrated sequence analysis (RISA) system--384-format
MEDLINE	sequencing pipeline with 384 multicapillary sequencer
PUBMED	Genome Res. 10 (11), 1757-1771 (2000)
REFERENCE	20530913
AUTHORS	4
TITLE	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kusui, P., Lewis, S., Matsuo, Y., Nakai, I., Pesole, G., Quackenbush, J., Schiml, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Buit, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Rung, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilmberg, L., Wyshaw-Borja, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S. and Hayashizaki, Y.
JOURNAL	Functional annotation of a full-length mouse cDNA collection
MEDLINE	Nature 409 (6821), 685-690 (2001)
PUBMED	21085660
REFERENCE	11217851
AUTHORS	5
TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
JOURNAL	Analysis of the mouse transcriptome based on functional annotation
MEDLINE	Nature 420, 563-573 (2002)
PUBMED	6 (bases 1 to 2682)
REFERENCE	6
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirokane, T., Horii, F., Imclani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraiki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
JOURNAL	Direct Submission
PUBMED	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

COMMENT

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

Location/Qualifiers

FEATURES

source

1..2682

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM.DB.9330209L13"

/db_xref="taxon:10090"

/clone="9330209L13"

/sex="male"

/tissue_type="dienecephalon"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

36..1001

/note="unannotated protein product; putative

sidoflexin 3 (MGDI|MG1:2137679, GB|NM_053197, evidence: BLASTN, 99%, match=2171)"

/codon_start=1

/protein_id="PAC28739.1"

/db_xref="GI:26330001"

/translation="MGDPLPLINIQEPRMDSITLGRARHEFTYTDPRNLISSEQL

ASRNIVNRYRAGVATPGLTEQILMARVYVDSAPHPTEGVVILGRASAVPNKMT

TGMLTFYRKYPTVFMQWQNSFNATVNSRSGDAPITVQDGTAVSATGAVAT

AAIGKSTFKHLPVGRFVPAVAANCILIPMRQRELOVGIPTVDEAGORLGHV

TAAGKGFVYVVISIGMAIPVMAIPVIMNTLEKDFLRPRPTGAPLVQVGFCLV

FAPPLCALAPQRSIHVTRFELPRRQIQANQSIDVYNNKGL"

2661..2666

/note="putative"

/note="putative"

2682

/note="putative"

CDS

BASE COUNT 565 a 763 c 626 g 728 t

ORIGIN

polyA_site

polyA_signal

2682

/note="putative"

/note="putative"

2682

/note="putative"

2682

Alignment Scores:

Pred. No.: 3.23e-140

Score: 1283.00

Percent Similarity: 98.08%

Best Local Similarity: 95.38%

Query Match: 95.39%

DB: 11

Gaps: 0

US-09-990-415a-2 (1-261) x AK034514 (1-2682)

OY 2 GlusertysmetgylguileuProleuAspIleasnIlegInguIupProArgrTPAspGln 21

|||||

DB 27 GAAGCAAGATGGGTGACCTGCTTAAATATCAACTCCAGAACCTCGGTGGACCAA 86

OY 22 SerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgrAsnLeuLeu 41

|||||

DB 87 AGCACAATTCAGGACAGACGCCGACATTCCTTCACAGTACTGATCCCGGAATCTGCTG 146

OY 42 LeuSerIlyAlaGlnLeuGluValAserArgAsnIleValGlnAsnIlyrArgAlaGlyVal 61

|||||

DB 147 CTGCTCGGGGAACGCTGGAAGCTTCCTCGGAACATCTGCGAATTCAGCGCTGCTG 206

OY 62 ValThrProGlyIleThrGluAspGlnLeuThrPargAlaIlyrValIlyrAspSerAla 81

|||||

DB 207 GCACACCCGGCTCATCTAGACACACCTATGGGACCCAAATACGTATGATCTCGCA 266

OY 82 PheHisProAspThrGlyGluIlyValIleValIleuIleGlyArgMetSerAlaGlnValPro 101

|||||

DB 267 TTCACATCCGACACGCGGGGAGAAAGTGTCTTGTATGCGGTATGTCACCCAGCTGCC 326

Ling Honglin the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library. "

Note: this is a NIH_MGC Library.

BASE COUNT	200 a	290 c	272 g	201 t
ORIGIN				

Alignment Scores:

Pred. No.:	6.1e-133	Length:	963
Score:	1216.00	Matches:	237
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	90.41%	Indels:	0
DB:	13	Gaps:	0

US-09-990-415A-2 (1-261) x BQ959114 (1-963)

Oy	102	Metsnmectrl1ietnlg1cysmetleuthrpherytrayglvsthrprothval1	121
Db	327	ATGACATGACCATTAATCGCTGCATGCTCACTTCTACAGAAAGACTCCAGCTGTGTC	386
Oy	122	PheTrpGlnTrpValaIngInSerPheAsnAlaIleValaAsnTyrSerAsnArgSerGly	141
Db	387	TTCTGGCAGTGGTCATCATCACTCTTCATATGCTATTGGATTACTCTAATCGCAGGGCC	446
Oy	142	AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThGlyAla	161
Db	447	GATCTCTCCCAACATGTGCACAGCACTTGGGGACACCATATGTAAGTGGCCACCATGGGGCT	506
Oy	162	ValAlaThrAlaLeuGlyLeuIleuSerLeuThrTyrSHisLeuProIleuValaGlyArg	181
Db	507	GTGGCTACTGCTCTGGGACTCAATCTCTCCACAGACCTGCCGCCGCTAGTGGTGTGA	566
Oy	182	PheValProPheAlaAlaValaAlaAlaAlaAsnCysIleAsnIleProIleuMetArgGln	201
Db	567	TTCTGTGCCCTTTGGACGTGTGGCCGCTGCCCACTGACATCAATCCCTCATGAGGCAG	626
Oy	202	ArgGluLeuGlnValaGlyIleProValaAlaAspGluIleAsnGlyIleArgLeuGlyTyrSer	221
Db	627	AGGAGGTGGCAGGGGGGCATCCAGTGACTATAGAGGTGGTGCAGAGGCTTGGGCACATCG	686
Oy	222	ValThrAlaAlaLysGlnGlyIlePheGlnValaValaIleSerArgIleCysMetAlaIle	241
Db	687	GTGACTGCTGCCAACAAGGAAATCTTCCAGGTGTGTGATATCAAGATCGGAATGGCAATT	746
Oy	242	ProAlaMetAlaIleProProIleuIleMetAspThrLeuGluLysAspPheLeuLys	261
Db	747	CCCCGCATGGCCATTTCCCCCGTGTATGTGAACACTGTGGAGAAAGAAAGACTTCTCGAAG	806
RESULT 10			
LOCUS	B0959114		
DEFINITION	AGENCOURT_10031240 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6482227		
ACCESSION	B0959114		
VERSION	B0959114.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 963)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: c9apbs.r@email.nih.gov		
	Tissue Procurement: DCTD/DMP		
	CDNA Library Preparation: Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	Plate: LNCM2664 row: n column: 20		
	High quality sequence stop: 635.		
FEATURES			
Source	Location/Qualifiers		
	1..963		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:6482227"		
	/tissue_type="carcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_id="NIH_MGC_40"		
	/note="Organ: prostate; Vector: pOTB7; site_1: XhoI;		
	site_2: EcoRI; cDNA made by oligo-dT priming.		
	directionally cloned into EcoRI/XhoI sites using the		
	following 5' adaptor: GCCACGAG(G). Library constructed by		

OY	25	leuGlytArgAlaArgGHisPhePheThrValThrAspProGlyAsnLeuLeuSerGly	44
Db	3	CTGGGAGAGCCCGGACCTTTTTCACGTCTTACTGATCTCCAAATCTCTCTGTCGGG	62
OY	45	AlaGlnLeuGlnAlaLeuSerArgAsnIleValGlnAsnTyrArgAlaGlyValValThrPro	64
Db	63	GCACAGCGTGAAGCTTCTCGGAACATCGTCGACAGAACTACAGGCCGCGGTGTGACCCCA	122
OY	65	GlyIleThrGluAspGlnLeuTyrArgValAlaSerTyrValTyrAspSerAlaPheHisPro	84
Db	123	GGGATCACCGAGGACCCAGCTGTGGAGGCCAAGTATGTGTAGTCTCCGCTTCATCCG	182
OY	85	AspThrGlyGluTyrValValLeuIleGlyArgMetSerAlaGlnValProMetAsnMet	104
Db	183	GACACAGGGAGAAAGGTGCTGATTTGGCCGATCGACGCCAGGTCGCCATGAACATG	242
OY	105	ThrIleThrGlyCysMetLeuThrPheTyrArgTyrThrProThrValValPheThrPro	124
Db	243	ACCATCACTGGCTCATGCTCATATTCATTCACAGGAAGACCCCAACCGTGTCTTGCGAG	302
OY	125	TyrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrPro	144
Db	303	TGGGTGATCATGCTCTTCAATGCCATTSTTAACTACTCCAAACCGAGTGGTGAACACTCC	362
OY	145	IleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThr	164
Db	363	ATCACTGTGAGGACGACGTGGGACAGCCATGTGATGATGCCACACTGGAGCTGTGGCCAG	422
OY	165	AlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValAlaArgPheValPro	184
Db	423	GCCTGGGACCTCAATCCCTACCAAGACCTGCCGCCCTTGTGTGGACGATTTGTGCC	482
OY	185	PheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGluLeu	204
Db	483	TTTTCAGACAGTGGAGCTGCCCACTGCATCAACATCCCCCGATGAGGCAAGAGAGCTG	542
OY	205	GlnValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyrSerValThrAla	224
Db	543	CAGGTGGGCATCCCGGTGCTGATGAGGACAGGTCAAGGCTTGGCTACTCGCTACTCA	602
OY	225	AlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMet	244
Db	603	GCCCAAGCAGGGAATCTTCCAGGTGTGATTTTCAAGAACTCTCATGGCATTCCTGCCATG	662
OY	245	AlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeuLys	264
Db	663	GCCATCCCACTGATCATGTGACACTCTGAGAGAAGAAAGACTTCCCTGAAG	713
RESULT 11			
BU506082			
LOCUS			
DEFINITION			
AGNCOURT_10018523 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6494119			
5', mRNA sequence.			
ACCESSION			
BU506082.1 GI:22812315			
KEYWORDS			
EST.			
934 bp mRNA linear EST 12-SEP-2002			

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 934)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLM14049 row: n column: 08
High quality sequence stop: 651.
Location/Qualifiers
1. 934
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6494119"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 198 a 265 c 268 g 202 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1,73e-132 Length: 934
Score: 1212.00 Matches: 244
Percent Similarity: 96.18% Conserved: 8
Best Local Similarity: 93.13% Mismatches: 8
Query Match: 90.11% Indels: 3
Gaps: 0
US-09-990-415a-2 (1-261) x BUS06082 (1-934)
QY 2 GlusertysmetgylguleuProleuasplleasnlleaglmguproarqtrpaspaln 21
Db 39 GAAGAGCAAGATGGGTGACTGCTTAAATATCAACATCCAGAACTCGGTGGACCAA 98
QY 22 SerThrPheleuglyaragalaarghlsphetherhrvalhraspproaarganleu 41
Db 99 AGCACATTTCTAGGCAAGCCGCGCATTTCTTCACTGACTGATCCCGAAATCTGCTG 158
QY 42 LeuserglyalaglnleuglualaserArgAsnillevalglnsenryrargalaglyal 61
Db 159 CTGTCGGGGAGAACGCTGGAAGCTCCCGGAACATCTCGAAGATTCACAGGCTGGTGTG 218
QY 62 ValThrProglylthethrgluaspGlnLeuTParqhlalasyrValTYraspserral 81
Db 219 GCAACCCCGGTCTCACTGAGACACAGCTATGGGACCAATACGTATGACTCGCA 278
QY 82 PheHisProasphrgrgluglylulysValValleuilegllyarqmetseralaglnvalpro 101
Db 279 TTCATCCGAGACGCGGGAGAGGTGCTTGTGTTGGCCGTATGTACGCCAGTCC 338
QY 102 MetAsnMetThrIleThrglyCysMetLeuThrPheTyraqlysThrProThrValVal 121
Db 339 ATGAACATGACATATAGCTGATGCTCACCCTTCACAGGAAGACTCCGACTGTGTG 398
QY 122 PheTrpIntrrpyalansglnserPheAsnAlalleValansnyrSerAsnArgSergly 141
Db 399 TTCGGCAGTGGTCATCATGCTCTCAATGCTATGTGATTAATCGACGGGC 458

QY 142 AspThrProIleThrValArgGlnleuglyThralatyValSerAlaThrThrGlyAla 161
Db 459 GATGCTCCCATCACTGTCGACGACTTGGGAGACACCTATGACTGGCACACGGGGCT 518
QY 162 ValAlaThrAlaLeuglyleulysSerleuThrlyshlsleuProProleuValGlyArg 181
Db 519 GTGGCTACTGCTCGGGAGCTCAAGTCTCCACCAAGACCTGCCCGCTAGTGTGCA 578
QY 182 PheValProPheAlaAlaValAlaAlaAlaAsnCytleasnlleProleuMetArgGln 201
Db 579 TTCCTGCTTTCGACGCTGTGGCCGCTGCCACAGTCAACATCCCTGATAGGACG 638
QY 202 ArgGlnleuglnValGlyIleProValAlaAspGlnlaaglylnargleuglyTyrSer 221
Db 639 AGGAGCTGACGAGGTGGGACATCCAGTACATGATGAGCTGTGACAGAGCTTGCCACTCG 698
QY 222 ValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIle 241
Db 699 GTGACTGCTCCCAA-CAGGAATCTTCAGGTGGATATCAAGAAATCGGAATGGCGAT 757
QY 241 eProAlaMetAlaIle-ProProleuIleMetAspThrLeuGluLysAspPheLeu 261
Db 758 TCCCGCCATGGCCATTCCTCCCGGTGATCATGAACATCTGGGAGAGAAATCTCTGA 817
QY 261 ys 261
Db 818 AG 819
RESULT 12
BM548674
LOCUS
DEFINITION
5', mRNA sequence.
ACCESSION
BM548674
VERSION
BM548674.1 GI:18783439
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1040)
NIH-MGC <http://mgi.nci.nih.gov/>
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLM12733 row: p column: 13
High quality sequence start: 16
High quality sequence stop: 631.
Location/Qualifiers
1. 1040
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5732220"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site:1: EcoRV (destroyed); Site:2: NotI; RNA source male hippocampus, age 27. Library is oligo-dt primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics

FEATURES

source

tracking code 012."

BASE COUNT 215 a 346 c 278 g 201 t

ORIGIN

Alignment Scores:

Pred. No.:	1,58e-130	Length:	1040
Score:	1196.00	Matches:	248
Percent Similarity:	96.89%	Conservative:	1
Best Local Similarity:	96.50%	Mismatches:	7
Query Match:	88.92%	Indels:	4
DB:	12	Gaps:	0

US-09-990-415a-2 (1-261) x BM548674 (1-1040)

QY 6 GYGLIuLeuProLeuaspIleasnIleGlnIuProArGTrpaspInserThrPheLeu 25

DB 3 GGGAAATTCCTTATACATCAACATCCAGGAACCTCGGGAGCAACAGT-ACCTTCCTG 61

QY 26 GYArGAlaArgHisPhePheThrValThrAspProArGAsnLeuLeuSerGlyAla 45

DB 62 GGCAGAGCCCGGACTTTTCACTGTTACTGATCCGAAATCTGCTGTCGGGGCA 121

QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValThrProGly 65

DB 122 CAGCTGGAGCTTCTCGAATCATCGTCAGAACTACAGGGCCGGCTGGTGCACCGAGG 181

QY 66 ILeThrGluAspGlnLeuTyrParGAlaLysTyrValTyrAspSerAlaPheHisProAsp 85

DB 182 ATCACCAGAGACAGCTGTGAGGGCCAGATATGTATGACTCCGCTTCATCCGGAC 241

QY 86 ThrGlyGluLysValIleuIleGlyArgMetSerAlaGlnValProMetAsnMetThr 105

DB 242 ACGGGGAGAGAGTGTCTCGATGGCCGATGTCAGCCAGGTGCCATGAACATGACC 301

QY 106 ILeThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrpGlnTrp 125

DB 302 ATCAGCTGGCTGATGCTCATCTTCACAGAAACCCCAACCTGGTGTCTGACAGTGG 361

QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThrProIle 145

DB 362 GTGAATCATCTTCAATGCCATGTTAACTACTCCAAACGCGAGTGTGACATCCCATC 421

QY 146 ThrValArgGlnLeuLysThrAlaTyrValSerAlaThrThcIlyAlaValAlaThrAla 165

DB 422 ACTGTGAGCAGCTGGGACAGCCTATGTAGTGCACACATGAGACTGTGGCCAGGCC 481

QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGlyArgPheValProPhe 185

DB 482 CTGGGACTCAATCCCTCACAAAGCACGCCCTTGTGTCGGCAGATTTTGCCCTTT 541

QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlnLeuGln 205

DB 542 GCAGAGCTGGCAGCTGCCAATGCATCAACATCCCTGATGAGCAGAGAGAGCTGAC 601

QY 206 ValGlyIleProValAlaAspGluAlaGlyAsnArgLeuGlyTyrSerValThrAlaAla 225

DB 602 GTGGGCAATCCCGGTGCTATGAGGAGGCTCAAAAGCTTGGCTACTCGGTGATGAGCC 661

QY 226 LysGlnGlyIlePheGlnValIleSerArgIleCysMetAlaIleProAlaMetAla 245

DB 662 AG-CAGGGAATCTTCAGAGGGGGATTC-AGAATGTGCATGGCCATTCCTGGCCAGGCC 719

QY 246 ILeProProLeuIleMetAspThrLeuGlu-LysLysAspPheLeuLys 261

DB 720 ATCCCCCTCTGATCATGACACTCTGGAAGAAAAAACTTCTTAAAA 768

RESULT 13

LOCUS BG775086 961 bp mRNA linear EST 15-MAY-2001

DEFINITION 602850054F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4760989 5',

ACCESSION BG775086

VERSION BG775086.1 GI:14045403

KEYWORDS EST. Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 961)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: gsaps-remail.nih.gov

Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Plate: LNCM612 row: 1 column: 14

High quality sequence stop: 765.

Location/Qualifiers

1. 961

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4760989"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_1ib="NIH_MGC_40"

/note="Organ: prostate; Vector: pOT7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(C). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 210 a 269 c 276 g 206 t

ORIGIN

Alignment Scores:

Pred. No.:	9.45e-130	Length:	961
Score:	1189.00 <td>Matches:</td> <td>236</td>	Matches:	236
Percent Similarity:	99.16% <td>Conservative:</td> <td>0</td>	Conservative:	0
Best Local Similarity:	99.16% <td>Mismatches:</td> <td>2</td>	Mismatches:	2
Query Match:	88.40% <td>Indels:</td> <td>1</td>	Indels:	1
DB:	12	Gaps:	0

US-09-990-415a-2 (1-261) x BG775086 (1-961)

QY 24 PheLeuGlyArgAlaArgHisPhePheThrValThrAspProArGAsnLeuLeuSer 43

DB 2 TTCCTGGGAGAGCCCGGCACTTTTCACTGTTACTGATCCGAAATCTGCTGCTGCC 61

QY 44 GYAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGlyValThr 63

DB 62 GGGGACAGCTGGAACTTCTCGAATCATGTCAGAACTCAAGAGGGCCGGCTGAC 121

QY 64 ProGlyIleThrGluAspGlnLeuTyrParGAlaLysTyrValTyrAspSerAlaPheHis 83

DB 122 CCGAGGATCACCGAGAACCGAGCTGTGAGGGCCCAAGATGTGTATGACTCGCTTCAT 181

QY 84 ProAspThrGlyLysValValLeuIleGlyArgMetSerAlaGlnValProMetAsn 103

DB 182 CCGAGACAGAGGAGAGAGGTCTGATGTGCCCATGTGCACCCAGGTGCCCATGAC 241

QY 104 MetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValValPheTrp 123

DB 242 ATGACCATCATCGGTGTCATGCTCATCTTCAAGAGAAAGCCCAACCGTGTGCTGG 301

QY 124 GlnTyrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGlyAspThr 143

DB 302 CAGTGGGTGAATGAGCTTCAATGCCATTTAACTACTACCAACCGCACTGTGACACT 361

```

QY 144 ProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAla 163
Db 362 CCCATCACTGTGAGCGAGCTGGGACAGCCATGTGATGTCACACACATGAGCTGTGGCC 421
QY 164 ThrAlaLeuGlyLeuLysSerIleuThrHisIleuProProLeuValGlyArgPheVal 183
Db 422 ACGGCCCTGGGACCTCAATTCCTCACCAAGCACTGGCCCCCTTGTGTGGCAGATTGTGTG 481
QY 184 ProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlu 203
Db 482 CCCTTGTGACGAGTGGGACGCTGCCACATGCATCAACATCCCTGATGATGAGCAGAGAGAG 541
QY 204 LeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThr 223
Db 542 CTGCAGGTGGGACATCCGGTGGCTGATGAGGACGAGCTCAGAGCTTGGCTACTCGTGACT 601
QY 224 AlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAla 243
Db 602 GCACCCCAAGCAGGAAATCTTCACAGTGTGTTCATTCAGAAATCTGCATTCGCAATTCCTGCC 661
QY 244 MetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeuLys 261
Db 662 ATGGACATCCCAACCATG-ATCATGAGACACTCTGGAGAAAGAACTTCTCTGAG 714

RESULT 14
BQ425697 862 bp mRNA linear EST 23-MAY-2002
LOCUS AGNCOURT_7887371 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6166944
DEFINITION 5', mRNA sequence.
ACCESSION BQ425697
VERSION BQ425697.1 GI:21121012
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 862)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@bgl.fda.gov
Tissue Procurement: ATCC/DCTD/DRP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM13528 row: b column: 01
High quality sequence stop: 603.
Location/Qualifiers
source 1. 862
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6166944"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 200 a 249 c 246 g 166 t 1 others
ORIGIN
Alignment Scores: 9.35e-129 Length: 862
Pred. No.: 1180.00 Matches: 234
Score: 94.82% Conservative: 4
Best Local Similarity: 93.23% Mismatches: 10

```

```

Query Match: 87.73% Indels: 3
DB: 13 Gaps: 2
US-09-990-415a-2 (1-261) x BQ425697 (1-862)

QY 1 MetGluSerLysMetGlyGlnLeuProLeuAspIleAsnIleGlnGlnProArgTrpAsp 20
Db 56 ATGGAAAGCAAAATGGGTGAATGGCTTTAACAATCAACATCCAGAACTCTGGTGGGAC 115
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValIleAspProArgAsnLeu 40
Db 116 CAAGTACTTCTCTGGGAGAGCCCGGACATTTTCACTGTACTGATTCCTCGAAATGTG 175
QY 41 LeuLeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
Db 176 CTGCTGTCCGGGACACACACTGGAAGCTTCGGAACATCGTGCACAACTACAGGCGCCGC 235
QY 61 ValValThrProGlyIleThrGlnAspGlnLeuThrPargAlaGlyTyrValTyrAspSer 80
Db 236 GTGGTGAACCCAGGAGATCCAGGAGACGAGCTGTGGAGGCCAAGTATGTATGACTTC 295
QY 81 AlaPheHisProAspThrGlyGlyLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db 296 GCCTTCCATCCGGACACAGGGAGAGAGGTGGTCTGATTTGGCCCATGTCACGCCAGGTG 355
QY 101 ProMetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db 356 CCCATGAACATGACCATCATGCTGCATGCTCATCATTCACAGGAAAGACCCCAACCGTGG 415
QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
Db 416 GTGTCTGGCAGTGGGTGAATCAGCTTCAATGACATTTGTAATCTACTCCAAACCGCAGT 475
QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThr-ThrG1 160
Db 476 GGTGACACTCCCAACACAGTGTAGGAGAGCTGGGAGACGCTATGTGAGTCCACACCGACTGG 535
QY 160 ValAlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValG1 180
Db 536 AGCTGTGGCCACGCGCCCTGGGACCTCAATTCCTCCCAAGACACGCGCCCTGTGGTGG 595
QY 180 ValArgPheValProPheAlaAlaValAlaAlaAsnCysIleAsnIleProLeuMetArg 200
Db 596 CAGATTGTGGCTCTTTCAGACAGTGGACAGCTGCCAATCTCAACATCCCTCGATATCA 655
QY 200 GlnArgGlnLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db 656 GGAGAGAGAGCTGCAGNGTGGGCAATCCCGTGGCTGATGAGGACAGTCAAGACTTGGCTA 715
QY 220 rSerValThrAlaAlaLysGlnGlyIlePheGln---ValValIleSerArgIleCysMe 239
Db 716 CTCGGTGACTCAGCAGCAAGCAGGAAATCTTCCAGAGGGGTGATTTCAAGAAATCTGCCA 775
QY 239 tAla---IleProAlaMetAlaIleProPro 248
Db 776 GGGCGGAATCCCGGCCCTGGCCCATCCCCC 806

RESULT 15
CA317398 731 bp mRNA linear EST 26-NOV-2002
LOCUS U1-M-FM0-cbn-g-19-0-UI_r1 NIH_BMAP_FM0 Mus musculus cDNA clone
DEFINITION IMAGE: 6811916 5', mRNA sequence.
ACCESSION CA317398
VERSION CA317398.1 GI:24535522
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 731)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished

```

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: egads-femail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 http://image.lnll.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES

Seq primer: pyx-5.
 Location/Qualifiers
 1..731

/organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE: 6811916"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH-BMAP-FW0"
 /note="Organ: Brain; Vector: pyx-Asc; Site:1: EcoR I;
 Site:2: Not I; The library was constructed according
 to Bernaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double stranded cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pyx-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGGAGAGACAG. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemlin Chin, Ph.D.,
 program coordinator."

BASE COUNT 163 a 201 c 206 g 159 t 2 others
 ORIGIN

Alignment Scores:

Pred. No.: 3,33e-127 Length: 731
 Score: 1166.00 Matches: 228
 Percent Similarity: 97.08% Conservative: 5
 Best Local Similarity: 95.00% Mismatches: 6
 Query Match: 86.69% Indels: 1
 DB: 14 Gaps: 0

US-09-990-415a-2 (1-261) x CA317398 (1-731)

QY 2 GUSerLySwetGlyLeuProLeuAspIleasnIleGlnGluProArgTrpAspGln 21
 |||||||
 DB 12 GAAGAGCAAGATGGTGACCTGCCCTTAATATCAACATCCAGACACCTGGTGGACCA 71
 QY 22 SerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 41
 |||||||
 DB 72 AGCACATTTCTAGAGAGAGCCCGCATTTCTTCACAGTCACTGATCCCCGAAATCTGCTG 131
 QY 42 LeuSerGlyAlaGlnLeuGlnIleAsnIleValGlnAsnTrpArgAlaGlyVal 61
 |||||||
 DB 132 CTGTCCGGGAGACAGCTGCAAGCTTCCCGACATGTCGAGATTCACAGGCTGGTGTG 191
 QY 62 ValThrProGlyIleThrGluAspGlnLeuTrpArgAlaAlaSerValTyrAspSerAla 81
 |||||||
 DB 192 GCMAACCCGGGTCTCACTGAGACAGCATATGCGAGCCAAATACGTATGACTCAGCA 251
 QY 82 PheHisProAspThrGlyGlyValValLeuIleGlyArgMetSerAlaGlnValPro 101
 |||||||
 DB 252 TTCCATCCCGGACACGGGGGAGAGGCTTGTATGGCCGTATGTCAGCCAGGTGCC 311
 QY 102 MetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgGlyThrProThrValVal 121
 |||||||

DB 312 ATGAACATGACCACTTACTGCTGCATGCTCCTTCTTACAGAGAACTCCGACTGTGCTG 371
 QY 122 PheTrpGlnTrpValAlaGlnSerPheAsnAlaIleValAsnTrpSerAsnArgSergly 141
 |||||||
 DB 372 TTTCGGCAGTGGGCAATCACTCTTCAATGCTATTGTGAATTTCTTAATCGACGGCC 431
 QY 142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAla 161
 |||||||
 DB 432 GATGCTCCCATCACTGTGACAGCAAGTGGGACAGCTATGTGAGTGCACACACTGGGGCT 491
 QY 162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArg 181
 |||||||
 DB 492 GTGCTACTGCTCTGGGACTCAAGTCTCTCAACAGCACTGCCCCGCTAGTGGTCGA 551
 QY 182 PheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeu-MetArgGly 201
 |||||||
 DB 552 TTGTGCTCTTGTGAGCTGTGGCCGCTGCGCAACTGCATCAACATCCCCCTTGATGAGGCA 611
 QY 201 nArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSe 221
 |||||||
 DB 612 GAGGAGACTGCAGGTGGGCAATCCAGTCACTGATGAGAGGCTGTGAGAGCTTGGCCACTC 671
 QY 221 rValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
 |||||||
 DB 672 GTGACTGTGTCACCAAGGGAATCTTCCAGGTGTATATCAAGAAATCGGAATGGCG 729

Search completed: October 16, 2003, 17:34:28
 Job time : 2083 secs

143

```

QY 349 GCTGTCGGCGTCACGCTGACCTCCCGCTGATGGTGGAGGGCCCGCGCAGAC 408
DB 146 GCTGTCGGCGTCACGCTGACCTCCCGCTGATGGTGGAGGGCCCGCGCAGAC 205
QY 409 GGAGGACAGAGAGAGAGGGGCTTTCAGAGCTTCAGAGAGCGATGGAAAGCAAAATGGG 468
DB 206 GGAGGACAGAGAGAGAGGGGCTTTCAGAGCTTCAGAGAGCGATGGAAAGCAAAATGGG 265
QY 469 AATTCCTTTAGACATCAACATCCAGAACCTCGCTGGAGCAAAAGTACTTCTTGGGCA 528
DB 266 AATTCCTTTAGACATCAACATCCAGAACCTCGCTGGAGCAAAAGTACTTCTTGGGCA 325
QY 529 GAGCGCGGCACTTTTACTGTTTCTGATCTGAAAATCTGCTCTGCGGGGACAGC 588
DB 326 GAGCGCGGCACTTTTACTGTTTCTGATCTGAAAATCTGCTCTGCGGGGACAGC 385
QY 589 TGAAGCTTCTCGGAACATGTCGAGAACATACAGAGGGCGCGCTGATCCCAAGATCA 648
DB 386 TGAAGCTTCTCGGAACATGTCGAGAACATACAGAGGGCGCGCTGATCCCAAGATCA 445
QY 649 CCGAGACACAGCTGTGAGAGGCCAAGTATGTATGACTCGGCTTCATCCGAGACAG 708
DB 446 CCGAGACACAGCTGTGAGAGGCCAAGTATGTATGACTCGGCTTCATCCGAGACAG 505
QY 709 GGGAGAAAGTGGCTGATGGCGCATGTCAGCCAGAGTGGCCATGAACATGACATCA 768
DB 506 GGGAGAAAGTGGCTGATGGCGCATGTCAGCCAGAGTGGCCATGAACATGACATCA 565
QY 769 CTGGCTGATGCTCAATCTACAGAGAACCCCAACCTGGTGTCTGCGAGTGGTGA 828
DB 566 CTGGCTGATGCTCAATCTACAGAGAACCCCAACCTGGTGTCTGCGAGTGGTGA 625
QY 829 ATCACTCTTCAATGCAATCTGTTAACTACTACACCGAGTGGTGAATCCCATCTG 888
DB 626 ATCACTCTTCAATGCAATCTGTTAACTACTACACCGAGTGGTGAATCCCATCTG 685
QY 889 TGAGGACAGCTGGGAGCAGCTGATGGAGTGGCCACCACTGAGAGTGGCCAGGCTG 948
DB 686 TGAGGACAGCTGGGAGCAGCTGATGGAGTGGCCACCACTGAGAGTGGCCAGGCTG 745
QY 949 GACTCAAAATCCCTCAACAGACAGCTGCGCCCTTGGTGGAGATTTTGGCTTTGAG 1008
DB 746 GACTCAAAATCCCTCAACAGACAGCTGCGCCCTTGGTGGAGATTTTGGCTTTGAG 805
QY 1009 CAGTGGGAGCTGGCCCAATGCTCAACATCCCTGATGAGGACAGAGAGTGGAGTGG 1068
DB 806 CAGTGGGAGCTGGCCCAATGCTCAACATCCCTGATGAGGACAGAGAGTGGAGTGG 865
QY 1069 GCATCCCGGTGGCTGATGAGGACAGTGGAGGCTTGGCTACTGCTGATGAGGAC 1128
DB 866 GCATCCCGGTGGCTGATGAGGACAGTGGAGGCTTGGCTACTGCTGATGAGGAC 925
QY 1129 AGGGAATCTTCCAGAGTGGATTTCAAGATCTGAGGAGTTCCTGAGGAGTTC 1188
DB 926 AGGGAATCTTCCAGAGTGGATTTCAAGATCTGAGGAGTTCCTGAGGAGTTC 985
QY 1189 CACCAGTATGATGAGACACTCTGAGAGAAAGAACTTCTGGAAG 1232
DB 986 CACCAGTATGATGAGACACTCTGAGAGAAAGAACTTCTGGAAG 1029

```

RESULT 2
US-08-905-223-165

; Sequence 165, Application US/08905223
; Patent No. 6222029

; GENERAL INFORMATION:

; APPLICANT: Edwards, Jean-Baptiste D.

; APPLICANT: Duclert, Aymeric

; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS

; NUMBER OF SEQUENCES: 503

; CORRESPONDENCE ADDRESS: Knobbe, Martens, Olson & Bear

```

STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 165:
SEQUENCE CHARACTERISTICS:
LENGTH: 315 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: s1g_peptide
LOCATION: 70..108
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 3.5
OTHER INFORMATION: seq MHLSTMNANPASS/RR
US-08-905-223-165

```

Query Match 5.3%; Score 65.8; DB 3; Length 315;
Best Local Similarity 56.1%; Pred. No. 1.2e-08;
Matches 124; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

```

QY 496 AACCTCGCTGGGACCAAGTACTTCTCGGCGCAGAGCCCGCACTTTTACTGTTACTG 555
DB 94 AACCGCGCTTCAGACAGACGCTCTTCTATGCGCGCTTCAGGCACTTCTTGATATCATCG 153
QY 556 ATCTCGCAAAATCTGCTGCTGTCGCGGCGCACAGCTGGAAGCTTCTCGGAACATGTCGAGA 615
DB 154 ACCCTCGCACACACTTGTGTACAGAGAGAGCTCTCAGAGAGGCTGTGACACTCTCTGAGG 213
QY 616 ACTAAGAGGCGCGCTGCTGATCCCAAGATCAACCGAGACAGTCTGTGGAGGCCAAGT 675
DB 214 ACTAAGAGGCGCGCTGCTGATCCCAAGATCAACCGAGACAGTCTGTGGAGGCCAAGT 716
QY 676 ATGTATATGACTCGCTCGCTTCATCCGAGACAGAGGAGAG 716
DB 274 AATCAAGCAGGCTATTCTACATCGGAGACACCAATGAGAG 314

```

RESULT 3

US-09-922-146-3

; Sequence 3, Application US/09922146
; Patent No. 656133

; GENERAL INFORMATION:

; APPLICANT: Lex M. Cowert

; APPLICANT: Brett P. Monia

; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 9 EXPRESSION

; CURRENT APPLICATION NUMBER: US/09/922,146

; NUMBER OF SEQ ID NOS: 48
; SEQ ID NO 3

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 16:00:17 ; Search time 77 Seconds
(without alignments)
1496.117 Million cell updates/sec

Title: US-09-990-415a-2

Perfect score: 1345
Sequence: 1 MESKMGELPLDINIOEPKWD.....PMAIPLIMDTLEKKDFLK 261

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool.P/US09990415/rnatc.16102003.073807.9464/app.query.fasta.1.455
-DB=Issued Patents.NA -QEXT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPT=0
-LOOPTXT=0 -UNITS=bits -STARF=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-OSER=US09990415 -RCGN_1_1_85_8turnat.16102003.073807.9464 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCT=100 -LONGLOG
-DEV_TIMEOUT=120 -NARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents.NA:*
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Dackfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	1345	100.0	1729	4	US-09-620-312D-1016
2	179	13.3	315	3	US-08-905-223-165
3	96	7.1	2943	1	US-07-923-976-3
4	92	6.8	2563	6	5422248-1
5	89.5	6.7	3024	4	US-07-923-976-7
6	89	6.6	1968	4	US-09-134-001C-2694
7	88	6.5	669	4	US-09-252-991A-1852
8	88	6.5	1185	4	US-09-252-991A-1931
9	85	6.3	552	4	US-09-252-991A-2024
10	84.5	6.1	2855	1	US-07-923-976-5
11	82	6.1	50341	1	US-08-247-901C-1
12	82	6.1	50341	2	US-09-075-904-1

13	82	6.1	52297	3	US-09-426-436-1	Sequence 1, Appl1
14	82	6.1	52297	4 <td>US-08-705-557-1</td> <td>Sequence 1, Appl1</td>	US-08-705-557-1	Sequence 1, Appl1
15	81	6.0	663	4 <td>US-09-252-991A-9294</td> <td>Sequence 9294, Ap</td>	US-09-252-991A-9294	Sequence 9294, Ap
16	81	6.0	918	4 <td>US-09-252-991A-9303</td> <td>Sequence 9303, Ap</td>	US-09-252-991A-9303	Sequence 9303, Ap
17	81	6.0	1428	4 <td>US-09-252-991A-9266</td> <td>Sequence 9266, Ap</td>	US-09-252-991A-9266	Sequence 9266, Ap
18	81	6.0	1545	4 <td>US-09-252-991A-9274</td> <td>Sequence 9274, Ap</td>	US-09-252-991A-9274	Sequence 9274, Ap
19	81	6.0	4403765	3 <td>US-09-103-840A-2</td> <td>Sequence 2, Appl1</td>	US-09-103-840A-2	Sequence 2, Appl1
20	81	6.0	4411529	3 <td>US-09-103-840A-1</td> <td>Sequence 1, Appl1</td>	US-09-103-840A-1	Sequence 1, Appl1
21	80.5	6.0	1153	4 <td>US-09-372-448A-5</td> <td>Sequence 5, Appl1</td>	US-09-372-448A-5	Sequence 5, Appl1
22	80	5.9	1866	4 <td>US-09-252-991A-3178</td> <td>Sequence 3178, Ap</td>	US-09-252-991A-3178	Sequence 3178, Ap
23	79.5	5.9	1884	4 <td>US-09-687-538B-5</td> <td>Sequence 5, Appl1</td>	US-09-687-538B-5	Sequence 5, Appl1
24	79.5	5.9	2237	1 <td>US-08-463-620-1</td> <td>Sequence 1, Appl1</td>	US-08-463-620-1	Sequence 1, Appl1
25	79.5	5.9	2237	2 <td>US-08-224-917-1</td> <td>Sequence 1, Appl1</td>	US-08-224-917-1	Sequence 1, Appl1
26	79.5	5.9	2237	5 <td>PCR-US95-03934A-1</td> <td>Sequence 1, Appl1</td>	PCR-US95-03934A-1	Sequence 1, Appl1
27	79.5	5.9	2237	5 <td>PCR-US95-03934A-1</td> <td>Sequence 1, Appl1</td>	PCR-US95-03934A-1	Sequence 1, Appl1
28	79.5	5.9	12412	1 <td>US-08-390-878-18</td> <td>Sequence 18, Appl1</td>	US-08-390-878-18	Sequence 18, Appl1
29	79.5	5.9	4411529	3 <td>US-09-103-840A-1</td> <td>Sequence 1, Appl1</td>	US-09-103-840A-1	Sequence 1, Appl1
30	79	5.9	1419	4 <td>US-09-252-991A-12594</td> <td>Sequence 12594, A</td>	US-09-252-991A-12594	Sequence 12594, A
31	79	5.9	2238	4 <td>US-09-252-991A-13219</td> <td>Sequence 13219, A</td>	US-09-252-991A-13219	Sequence 13219, A
32	78.5	5.8	2340	4 <td>US-09-786-240-29</td> <td>Sequence 29, Appl1</td>	US-09-786-240-29	Sequence 29, Appl1
33	78.5	5.8	2350	4 <td>US-09-620-312D-1094</td> <td>Sequence 1094, Ap</td>	US-09-620-312D-1094	Sequence 1094, Ap
34	78.5	5.8	9603	4 <td>US-09-203-895-1</td> <td>Sequence 1, Appl1</td>	US-09-203-895-1	Sequence 1, Appl1
35	78.5	5.8	4403765	3 <td>US-09-103-840A-2</td> <td>Sequence 2, Appl1</td>	US-09-103-840A-2	Sequence 2, Appl1
36	78	5.8	2250	4 <td>US-09-221-017B-1033</td> <td>Sequence 1033, Ap</td>	US-09-221-017B-1033	Sequence 1033, Ap
37	77.5	5.8	2028	4 <td>US-09-221-017B-1001</td> <td>Sequence 1001, Ap</td>	US-09-221-017B-1001	Sequence 1001, Ap
38	77.5	5.8	6027	4 <td>US-09-620-312D-517</td> <td>Sequence 517, Ap</td>	US-09-620-312D-517	Sequence 517, Ap
39	77.5	5.8	25165	4 <td>US-09-453-702B-39</td> <td>Sequence 39, Appl</td>	US-09-453-702B-39	Sequence 39, Appl
40	77	5.7	864	4 <td>US-09-252-991A-1297</td> <td>Sequence 1297, Ap</td>	US-09-252-991A-1297	Sequence 1297, Ap
41	77	5.7	1050	4 <td>US-09-252-991A-1269</td> <td>Sequence 1269, Ap</td>	US-09-252-991A-1269	Sequence 1269, Ap
42	77	5.7	1308	4 <td>US-09-252-991A-1179</td> <td>Sequence 1179, Ap</td>	US-09-252-991A-1179	Sequence 1179, Ap
43	77	5.7	1878	3 <td>US-08-996-139-14</td> <td>Sequence 14, Appl</td>	US-08-996-139-14	Sequence 14, Appl
44	77	5.7	1878	3 <td>US-08-995-659-14</td> <td>Sequence 14, Appl</td>	US-08-995-659-14	Sequence 14, Appl
45	77	5.7	1878	3 <td>US-09-215-649A-14</td> <td>Sequence 14, Appl</td>	US-09-215-649A-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-620-312D-1016
Sequence 1016, Application US/09620312D
Patent No. 6569662

GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyang
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunging
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhilwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 1016
LENGTH: 1729
TYPE: DNA

ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (247)..(1371)
 US-09-620-312D-1016

Alignment Scores:

Pred. No.:	5,72e-165	Length:	1729
Score:	1345.00	Matches:	261
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-990-415a-2 (1-261) x US-09-620-312D-1016 (1-1729)

```

QY 1 MetGluSerIysMetGlyIleuProLeuAspIleuSnlleGlnIupProArGTrpAsp 20
DB 247 ATGGAAGCAAAATGGGTGAATGCTTTAGACATCAACATCCAGAACCTCGCTGGGAC 306
QY 21 GluSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArGTrpAsn 40
DB 307 CAAAGTACTTTCCTGGGACAGAGCCCGGACATTTTCACTGTACTGATCCGCAATCTG 366
QY 41 LeuLeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTrpArgAlaGly 60
DB 367 CTGCTGTCGGGGGACAGCTGAGAGCTTCTCGAACAATCTCGAACAATACAGGGCCGGC 426
QY 61 ValValThrProGlyIleThrGlnAspGlnLeuTrpArgAlaGlyValTrpAspSer 80
DB 427 GTGTGAGCCCGAGGATCACCGAGAGCCAGCTGTGGAGGGCCAAATATGTATGACTCC 486
QY 81 AlaPheHisProAspThrGlyIleValValIleuIleGlyArgMetSerAlaGlnVal 100
DB 487 GCCTTCCTCCGGGACACAGGGGAGAGAGTGGTCCGATTTGGCCGATGTACAGCCAGGTG 546
QY 101 ProMetAsnMetThrIleThrGlyCysMetLeuThrPheTrpArgIleValThrProThr 120
DB 547 CCCATGAAACATGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606
QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTrpSerAsnArgSer 140
DB 607 GTGTCTGCACTGGGTGATATCATCTTCAATGCCATTGTACTACCCCAACGGCAGT 666
QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTrpValSerAlaThrThrGly 160
DB 667 GGTGACACCCCATCATCTGTAGGACAGCTGGGAGACAGCTATGTGTGCCACCACTGGA 726
QY 161 AlaValAlaThrAlaLeuGlyLeuIysSerLeuThrIleHisLeuProProIleuValGly 180
DB 727 GCTGTGGCCACGCGCTGGGACTCAATCTCCACACCCACCTGCCCCCTGTGGTGGC 786
QY 181 ArgPheValProPheAlaValAlaValAlaAsnCysIleAsnIleProLeuMetArg 200
DB 787 AATTTTGTCCCTTGTGACAGAGTGGCAGCTGCCAATGCATCAACATCCCCCGTGAAGG 846
QY 201 GlnArgGlnLeuGlnValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTrp 220
DB 847 CAGAGAGACTGAGAGTGGGACATCCGCGTGGTGAAGAGGACAGTCAAGAGCTTGGCTAC 906
QY 221 SerValThrAlaAlaIysGlnGlyIlePheGlnValAlaIleSerArgIleCysMetAla 240
DB 907 TCGGTACGTGACACCAAGCAGGGAATCTTCCAGTGGTGTATTTCAAGAAATCTCAATGGG 966
QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnIuIysAspPheLeu 260
DB 967 ATTCCTGCAATGGGCATCCACCACTGATATGACACTCTGGAGAAAGAAAGACTTCTCTG 1026
QY 261 Lys 261
DB 1027 AAG 1029
RESULT 2
  
```

US-08-905-223-165

Sequence 165, Application US/08905223

Patent No. 6222029

GENERAL INFORMATION:

APPLICANT: Edwards, Jean-Baptiste D.

APPLICANT: Ducloux, Aymeric

TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS

NUMBER OF SEQUENCES: 503

CORRESPONDENCE ADDRESS:

ADDRESS: Knodde, Martens, Olson & Bear

STREET: 501 West Broadway

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92101-3505

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Win95

SOFTWARE: Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/905,223

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Isrelsen, Ned A.

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 165:

SEQUENCE CHARACTERISTICS:

LENGTH: 315 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: DOUBLE

TOPOLOGY: LINEAR

MOLECULE TYPE: CDNA

ORGANISM: Homo sapiens

TISSUE TYPE: Brain

FEATURE:

NAME/KEY: sig-peptide

LOCATION: 70..108

IDENTIFICATION METHOD: Von Heijne matrix

OTHER INFORMATION: score 3.5

US-08-905-223-165

Alignment Scores:

Pred. No.:	2,62e-14	Length:	315
Score:	179.00	Matches:	29
Percent Similarity:	66.22%	Conservative:	20
Best Local Similarity:	39.19%	Mismatches:	25
Query Match:	13.31%	Indels:	0
DB:	3	Gaps:	0

US-09-990-415a-2 (1-261) x US-08-905-223-165 (1-315)

```

QY 16 GluProArgTrpAspGlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThr 35
DB 93 AAMCCCGGCTTCAGAGAGCTCTATGCGCGCTTCAGGCACTTGTGATATCATC 152
QY 36 AspProArgAsnLeuLeuLeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGln 55
DB 153 GACCCCTGCGACACTCTTGTCACTGAGAGACGTCTCAGAGAGCTGTGACGTGTGAG 212
QY 56 AsnTrpArgAlaGlyValValThrProGlyIleThrGlnAspGlnLeuTrpArgAlaLys 75
DB 213 GACTATTAAGCATGGACCCCTGCGCGGGGTACCAATGACAGCTCTGTGAGTGCACAG 272
QY 76 TyrValTrpAspSerAlaPheHisProAspThrGlyIuLys 89
  
```

```

Db      273 AAAATCAAGCAGGCTATTCTACATCCGACACACATGAGAAG 314
RESULT 3
US-07-923-976-3
: Sequence 3, Application US/07923976
: Patent No. 5574136
: GENERAL INFORMATION:
:   APPLICANT: Nagata, Shigekazu
:   APPLICANT: Fukunaga, Rikio
:   TITLE OF INVENTION: DNA Encoding Granulocyte
:   NUMBER OF SEQUENCES: 8
:   ADDRESS/SEE: Jones, Tullar & Cooper, P.C.
:   STREET: P.O. Box 2266 Eads Station
:   CITY: Arlington
:   STATE: Virginia
:   ZIP: 22202
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: Patent Release #1.0, Version #1.25
:   CURRENT APPLICATION DATA:
:     APPLICATION NUMBER: US/07/923, 976
:     FILING DATE: 19920922
:     CLASSIFICATION: 435
:     PRIOR APPLICATION DATA:
:       APPLICATION NUMBER: JP 74539/1990
:       FILING DATE: 23-MAR-1990
:       PRIOR APPLICATION DATA:
:         APPLICATION NUMBER: JP 176629/1990
:         FILING DATE: 03-JUL-1990
:         PRIOR APPLICATION DATA:
:           APPLICATION NUMBER: PCT/JP91/00375
:           FILING DATE: 22-MAR-1991
:           ATTORNEY/AGENT INFORMATION:
:             NAME: Hellwege, James W.
:             REGISTRATION NUMBER: 28,808
:             REFERENCE/DOCKET NUMBER: 514853
:             TELECOMMUNICATION INFORMATION:
:               TELEPHONE: 703-415-1500
:               TELEFAX: 703-415-1508
:             INFORMATION FOR SEQ ID NO: 3:
:               SEQUENCE CHARACTERISTICS:
:                 LENGTH: 2943 base pairs
:                 TYPE: NUCLEIC ACID
:                 STRANDEDNESS: single
:                 TOPOLOGY: linear
:               MOLECULE TYPE: cDNA to mRNA
:               FEATURE:
:                 NAME/KEY: CDS
:                 LOCATION: 170..2677
:                 US-07-923-976-3
:
: Alignment Scores:
: Pred. No.: 0.0693      Length: 2943
: Score: 96.00      Matches: 59
: Percent Similarity: 42.86%      Conservative: 34
: Best Local Similarity: 27.19%      Mismatches: 89
: Query Match: 7.14%      Indels: 36
: DB: 1      Gaps: 12
:
US-09-990-415a-2 (1-261) x US-07-923-976-3 (1-2943)
QY      61 ValValIhPrroGlyIleThrGluAspGlnLeuTrpArgAlaLys-----TyrValTyr 78
DB      1676 ATCGTACTGCC--TTGTACAGGACACATGACCCCTCCACATGCTATGCTAC 1732
QY      79 AspsrAlaPhelIsProAspThrGlyGluLysValIleu-----IleGlyArgMet 96
DB      1733 TCTCAAGAAATGGCTCCCTCCATGCCACAGAGCTGATCTTAAGCACATTTGGCAAGACC 1792

```

```

QY      97 SerAlaGln-----ValProMetAsnMetThrIleThrGlyCysMetLeuThrPhe 113
DB      1793 TGGCCACAGCTGAGTGGGTGCTGAGCCCTGAGCTGGGGAAGAGCCCTTACCCAC 1852
QY      114 TyrArgLysThrProThrValValPheTrp---GlnTTPValAsnGlnSerPheAsnAla 132
DB      1853 TAC-----ACCATCTTCTGGACCAACGCTCAACACAGCTCTTCTCCGCC 1897
QY      133 IleValAsnTyrSerAsnArgSer-----GlyAspThrProIleThr----- 146
DB      1898 ATCCGAATGGCTCTCCCGTGGCTTTGTCTTCATGCGCTGGAGCCGCCAGTGTGAT 1957
QY      147 ---ValArgGlnLeuGlyThrAlaTyrValSerAlaThrIleThrGlyAlaValAlaThrAla 165
DB      1958 CACATCCACATCATGTGCTGCGACGAGCTGGGGCCACCAACAGTACAGTCTCACCCG 2017
QY      166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValPro--- 184
DB      2018 ATGACCTTGACCCCGACAGGGGTGCGAGCTACACATCATCTGGCGCTGTCG-CTCTCT 2076
QY      185 ---PheAlaAlaValAlaAlaAlaAsnGlyIleAsnIleProLeuMetArgGlnArgGlu 203
DB      2077 GCTGTGCTCACCTGCTCTGTGAGTGCCTGGCTG---TTGACGCCCAACAGAGA 2133
QY      204 LeuGln-ValGlyIleProValAlaAspGlnAlaGlyGlnArgLeuGlyTyrSerVal-- 222
DB      2134 GAATCCCTCTGCGCCACAGTGTGCCAGACCCAGCTCACAGACCTGCGGTCTCGGTGCC 2193
QY      223 -ThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIlePr 242
DB      2194 CACAAATCATGGAGGAGATGCTTCAG-----CTGCC 2226
QY      242 cAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysAsp 258
DB      2227 CGGCTTGACAGCCACCCATCACCAGCTCAGTGTGAGGAGAT 2275
RESULT 4
5422248-1
: Patent No. 5422248
: APPLICANT: SMITH, CRAIG A.; LARSEN, ALF D.; SIMS, JOHN E.;
: BENSON, CURTIS M.
: TITLE OF INVENTION: DNA SEQUENCES ENCODING GRANULOCYTE-COLONY
: STIMULATING FACTOR RECEPTORS
: NUMBER OF SEQUENCES: 6
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/08/6,183
:   FILING DATE: 15-JAN-1993
:   PRIOR APPLICATION DATA:
:     APPLICATION NUMBER: 587,329
:     FILING DATE: 24-SEP-1990
:     APPLICATION NUMBER: 522,952
:     FILING DATE: 03-APR-1990
:     APPLICATION NUMBER: 416,306
:     FILING DATE: 03-OCT-1989
:     APPLICATION NUMBER: 412,816
:     FILING DATE: 26-SEP-1989
:   SEQ ID NO: 1:
:     LENGTH: 2503
:     5422248-1
:
: Alignment Scores:
: Pred. No.: 0.183      Length: 2563
: Score: 92.00      Matches: 58
: Percent Similarity: 42.86%      Conservative: 35
: Best Local Similarity: 26.73%      Mismatches: 89
: Query Match: 6.84%      Indels: 36
: DB: 6      Gaps: 12
:
US-09-990-415a-2 (1-261) x 5422248-1 (1-2563)
QY      61 ValValIhPrroGlyIleThrGluAspGlnLeuTrpArgAlaLys-----TyrValTyr 78

```



```

Db      1688 ATCGTAGCTCC---TTGTACAGAGAACACCATGGAGCCCTCCAGCATGCTATGCTTAC 1744
QY      79 AspSerAlaPheHisProAspThrGlyGluValValLeu-----IleGlyArgmet 96
Db      1745 TCTCAAGAAATGGCTCCCTCCAGAGCCCGCCAGAGCTATGTAAGACATTTGGCAAGACC 1804
QY      97 SerAlaGln-----ValProMetAsnMetThrIleThrGlyLysMetLeuThrPhe 113
Db      1805 TGGGCACAGCTGAGAGTGGGTGCTTGAGCCCTGAGCTGGGGAGAGCCCTTACCAC 1864
QY      114 TyrArgLysThrProThrValAlaPheTrp---GlnTrpValAsnGlnSerPheAsnAla 132
Db      1865 TAC-----ACCATCTCTGAGACCAAGCGTCAGAACCAAGTCCTTCTCGGCC 1909
QY      133 IleValAsnTrpSerAsnArgSer-----GlyAspThrProIleThr----- 146
Db      1910 ATCCCTGAATGCCCTCCCTCCGCTGCTTGTCTCCATGAGCTGGAGCCCGCCAGCTGTAT 1969
QY      147 ---ValArgGlnLeuGlyThrAlaThrValSerAlaThrGlyAlaValAlaThrAla 165
Db      1970 CACATCCACCTCATGGCTCCAGCCAGGCTGGGGCCACCAACAGTACAGTCTCTCACCTG 2029
QY      166 LeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGlyArgPheValPro--- 184
Db      2030 ATGACCTTGACCCCGACAGGGGGTCCGAGCTACACATCATCTCTGGCTGTCG-CTTCT 2088
QY      185 ---PheAlaIleValAlaAlaAlaAsnGlyIleAsnIleProLeuMetArgGlnArgGlu 203
Db      2089 GCTGTGGCTACCTGCTCTGTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2145
QY      204 LeuGln-ValGlyIleProValAlaAspIleValGlnArgLeuGlyIleTyrSerVal-- 222
Db      2146 GAATCCCTCTGCGCCAGTGTCCCAAGACCCAGCTCAGACAGCTGGCTGCTGCTGCTGCTG 2205
QY      223 -ThrAlaAlaLysGlnGlyIlePheGlnValAlaIleSerArgIleCysMetAlaIlePr 242
Db      2206 CACAATCATGAGAGAGATGCTTCACG-----CTGCC 2238
QY      242 AlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAsp 238
Db      2239 GCGCTTGGCAGCCAGCCACCATCACCAGCTCAGAGTGTGAGAGAGAT 2287

RESULT 5
US-07-923-976-7
; Sequence 7, Application US/07923976
; Patent No. 5574136
; GENERAL INFORMATION:
; APPLICANT: Nagata, Shigekazu
; APPLICANT: Fukunaga, Rikio
; TITLE OF INVENTION: DNA Encoding Granulocyte
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones, Tullar & Cooper, P.C.
; STREET: P.O. Box 2266 Eads Station
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,976
; FILING DATE: 19920922
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 74539/1990
; FILING DATE: 23-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 176629/1990
; FILING DATE: 03-JUL-1990

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP91/00375
; FILING DATE: 22-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hellwege, James W.
; REGISTRATION NUMBER: 28,808
; REFERENCE/DOCKET NUMBER: 514853
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-415-1500
; TELEFAX: 703-415-1508
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3024 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 170..2758
; US-07-923-976-7

Alignment Scores:
Pred. No.: 0.509 Length: 3024
Score: 89.50 Matches: 55
Percent Similarity: 36.73% Conservative: 35
Best Local Similarity: 22.45% Mismatches: 90
Query Match: 6.65% Indels: 65
DB: 1 Gaps: 13

US-09-990-415a-2 (1-261) x US-07-923-976-7 (1-3024)
QY      61 ValValaThrProGlyIleThrGlnAspGlnLeuTrpArgAlaLys-----TyrValTyr 78
Db      1676 ATCGTAGCTCC---TTGTACAGAGAACACCATGGAGCCCTCCAGCATGCTATGCTTAC 1732
QY      79 AspSerAlaPheHisProAspThrGlyGluValValLeu-----IleGlyArgmet 96
Db      1733 TCTCAAGAAATGGCTCCCTCCAGAGCTCCAGAGCTGACATCTAAGCACATTTGGCAAGACC 1792
QY      97 SerAlaGln-----ValProMetAsnMetThrIleThrGlyLysMetLeuThrPhe 113
Db      1793 TGGGCACAGCTGAGAGTGGGTGCTTGAGCCCTGAGCTGGGAAGAGCCCTTACCAC 1852
QY      114 TyrArgLysThrProThrValAlaPheTrp---GlnTrpValAsnGlnSerPheAsnAla 132
Db      1853 TAC-----ACCATCTCTGAGACCAAGCGTCAGAACCAAGTCCTTCTCGGCC 1897
QY      133 IleValAsnTrpSerAsnArgSer-----GlyAspThrProIleThr----- 146
Db      1898 ATCCGAAATGGCTCCCTCCGCTGCTTGTCTCCATGAGCTGGAGCCCGCCAGTCTGTAT 1957
QY      147 ---ValArgGlnLeuGlyThrAlaThrValSerAlaThrGlyAlaValAlaThrAla 165
Db      1958 CACATCCACCTCATGGCTGCTCCAGAGCTGGGGCCACCAACAGTACACTCTCCACCTG 2017
QY      166 LeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGlyArgPheValProPhe 185
Db      2018 ATGACCTTGACCCCGACAGGGGTCCGAGCTACACATCATCTCTGGCCCTG-----TTC 2068
QY      186 AlaAlaValAlaAlaAlaAsnGlyIle----- 194
Db      2069 GCGCTCTGCTGCTGCTCAGCTGCTGTGAGACTGCTGCTGCTGTGTCAGCCCAAC 2128
QY      195 ---AsnIleProLeuMetArgGlnArgGluLeuGlnValGly 207
Db      2129 AGGAGAAATCCCTCTGCGCCAGTGTCCAGAACCCAGCTCAGACAGCTGGGTCTCTG 2188
QY      208 IleProValAlaAspGlu-----AlaGlyGlnArgLeuGlyTyrSer 221
Db      2189 GTGCCACATCATGAGAGAGCTGCCGAGCCAGACAGAGGACAGTGGCTGGGGAGACA 2248
QY      222 ValThrAlaAlaLys-----GlnGlyIlePheGlnValVal 233

```

```

Db      2249 TGTGAATGAGCGCTCTCAACCCACATCTTGTGTGATGCCATCCCTCCAG----- 2302
QY      234  ILeSerArgIleCysMetAlaIleProAlaMetAlaIleProProIleuMetAspThr 253
Db      2303 -----CTGCCCGGCTTGGCAGCGCCACCCATCACCAAGCTCACA 2341
QY      254  LeuGIuLysLysAsp 258
Db      2342 GTGCTGAGAGAGAT 2356

RESULT 6
US-09-134-001C-2694
; Sequence 2694, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2694
; LENGTH: 1968
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2694

Alignment Scores:
Pred. No.:      0.29      Length:      1968
Score:          89.00     Matches:      54
Percent Similarity: 33.338  Conservative: 32
Best Local Similarity: 20.938  Mismatches: 78
Query Match:    6.62%      Indels:      94
DB:            4          Gaps:      10

US-09-990-415a-2 (1-261) x US-09-134-001C-2694 (1-1968)
QY      53  ILevalGlnAsnTyrArgAlaGlyValAlaThrProGlyIleThrGlnAspGlnLeuTrp 72
Db      1219 ATTATGCTGCTGTTCTTACAGTATTTACTCAAGAAAT----- 1260
QY      73  ArgAlaLysTyrValTyrIspSerAlaPheHisProAspThrGlyLysValValLeu 92
Db      1261 -----AAATATATTATACGTAAATTCACAAAGCAATTGAAGCGTTGAACCTACTCTTA 1314
QY      93  ILeGlyArgMetSerAlaGlnValProMetAsnMetThrIleThrGlyCysMetLeuThr 112
Db      1315 ATCTATCCATTA-----CTTCAGTTTCATTAATCTGATGTTTAATGTCG 1359
QY      113 PheTyrArgLysThrProThrValValPheTyrGlnTrpValAsnGlnSerPheAsnAla 132
Db      1360 TATGCTTTAATCCGACACACT-----TGCTTAAT-----CATTTG 1398
QY      133  ILeValAsnTyrSerAsnArgSerGlyAspThrProIleThrValArgGlnLeu----- 150
Db      1399 CTACTAATATGTTAAATAGTCTGTACAGGCTCTAATATATGTTATGTTGGTCTTGTGATT 1458
QY      151 -----GlyThrAlaTyrValSerAla 157
Db      1459 GGAGCTATGATGGCAATTACATGGCGGTCATTTAATAAGACGCTTATGTATTGCT 1518
QY      158 ThrThrGly----- 160
Db      1519 ACTGCTGATTAAGAGAAATGACAGCTCCTATTACTGACAGCATGATGGTGGTATG 1578
QY      161 -----AlaValAlaThrAla----- 165

```

```

Db      1579 ATTCTCATTTACTATCGCTACAGCAATGTTAATCTTTAGAAAGAAATTCATAAGAG 1638
QY      166 -----LeuGlyLeuIspSerLeuThrLysHisLeu 175
Db      1639 CAAAGAGCTCAATTCTTACCAACTATGTCATGGATATATCATTTTACTGAAGTGCC 1698
QY      176  ProProLeuValGlyArgPheValProPheAlaAlaValAlaAlaAsnCysIleAsn 195
Db      1699 -----ATCCATTCGACGCTGCCAT----- 1719
QY      196  ILeProLeuMetArgGlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGly 215
Db      1720 ---CCATTACGTGTGATTCCTTCATCAATGATGGCTGCTGACGTGTGAGCAATGCT 1776
QY      216  GlnArgLeuGlyTyrSerValThrAlaAlaLysGlnGlyIlePheGlnValAla----- 233
Db      1777 TTAGGTTTAGATCAAGTATCAAGACCACTCATGATGGAAATTTCTGCTATTATGGTACT 1836
QY      234  ---ILeSerArgIleCysMetAlaIleProAlaMetAlaIleProProLeuIle 250
Db      1837 GATTTTAACCATTACTTCAAACTCTTATGTTGTGGGTACGTTAGTT 1890

RESULT 7
US-09-252-991A-1852
; Sequence 1852, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONA
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1852
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1852

Alignment Scores:
Pred. No.:      0.0654     Length:      669
Score:          86.00     Matches:      46
Percent Similarity: 34.968  Conservative: 40
Best Local Similarity: 18.70%  Mismatches: 64
Query Match:    6.54%      Indels:      96
DB:            4          Gaps:      11

US-09-990-415a-2 (1-261) x US-09-252-991A-1852 (1-669)
QY      35  ThrAspProArgAsnLeuLeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleVal 54
Db      95  AGCAGACCCGCTGTCGATGCGACATCGACACCATGCTCTACCCCAATACCCCAACTACGTC 154
QY      55  GlnAsnTyrArgAlaGlyValAlaThrProGlyIleThrGlnAspGlnLeuTrpAlaGly 74
Db      155  GAGAACTACGCGCGCGCGCTGTC-----CAGTTCTGGCGCGCC 193
QY      75  LysTyrValTyrAspSerAla----- 81
Db      194  GGG-----CTGATATTCGCGCGCGCGCGCAAGCCGTCATGCTGAGCGCGCTACTGG 247
QY      82  PheHisProAspThrGlyLysValValIleuIleGlyArgMetSerAlaGlnValPro 101
Db      248  TTCACCCCGAGCTGGAAGACGCGCGCGCATGAT-----CCGCGCGCATGCGCC 298
QY      102  MetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValVal 121
Db      299  ATGTCATGATGACATATCGGACCATGCTG-----ACTGCATGCTGCTG 343

```

QY	122	TherprgIntPrvAlaAsnGlnSerPheAsnAlaIleValAsnTYrSerAsnArgSeryl	141
Db	344	GCGCGGATGGGACGCGCGCATGGAGGACAGTGTCTCC	385
QY	142	AspThrProIleThhValArgGlnLeuGlyThrAlaTYrValSerAlaThhThcGlyAla	161
Db	386	--ACGCCAGCCTCGCTGGCCGAAT	409
QY	162	ValAlaThrAlaLeuGlyLeuLYrSerLeuThrLYrSHsLeuProLeuValGlyArg	181
Db	410	-----CTATCGCAG	421
QY	182	PheValProPheAlaAlaVal	194
			AlaAlaAlaAsnCysIle
Db	422	CTGCTGCGGTACTTCTGACTGCGCATGCTGTCCAGCTGGGCGCGCTGCGGATG	481
QY	195	AsnIleProLeuMetArgGlnArgGlnLeuGlnValGlyIleProValAlaAspGlnAla	214
Db	482	TTTCG	502
QY	215	GlyGlnArgLeuGlyTYrSerValThhAlaAlaAsnGlnGlyIlePheGlnValIle	234
Db	503	-----GCGGCTCTTGTCTGCTCCCTGTGTC	529
QY	235	SerArgIleCysMetAlaIleProAlaMetAlaIleProLeuIleMetAspThhLeu	254
Db	530	TTGTCGCGCGGTCTTCATCGTCCCGCGCTGGCT	586
QY	255	GluLYrSlyAspPheLeu	260
Db	587	GCGGCAACCAAGTCTCTG	604

```

RESULT 8
US-09-252-991A-1931
; Sequence 1931, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERGININOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1931
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1931

Alignment Scores:
Pred. No.:      0.169           length:       1185
Score:          88.00         Matches:        46
Percent Similarity: 34.96%     Conservative:   40
Best Local Similarity: 18.70%    Mismatches:    64
Query Match:     6.54%         Indels:        96
DB:              4            Gaps:          11

US-09-990-415A-2 (1-261) x US-09-252-991A-1931 (1-1185)

QY      35 ThrAsPProArGAnsLeuleLeuSerGIyAlagInleuGluaIsErArGsAnIlleaI 54
      : :::::||||| : :::::~::~ ~:: |||| |~~~
Db      415 AGCGAGCCGCtGTGGTCAGATCGTACACCGAGTGCTCCTACCAATAcCGCCAaCTACTGTC 474

QY      55 GlAsnTrArGaLagIyValValThPrOgIIeThrGLuspgInleuTIrPaRaLa 74
      : ::||~~~~~| |~~~~~| |~~~~~| |~~~~~| |~~~~~|
Db      475 GAgaActAcACccCGcgGgcGtgGtc-----CAGTTCTGggCGGCC 513

```

```

QY      75 LysTyrValTyrAspSerAla----- 81
Db      514 GGG-----CTGATGTCTCGCGCGCCGCGCAAGCCGTCTGAGCCGCGCTACTGG 567
QY      82 PheHisProAspThrGlyGluValValIleuIleGlyArgMetSerAlaGlnValPro 101
Db      568 TTCACACCCGAGACTGGAAAGCCGCGCCGCGCACTATT-----CCGCGCGCCATCGCC 618
QY      102 MetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 121
Db      619 ATCGTCATGACCATCATCTACGCAACATGCTG-----ACTGCATGTGTGTG 663
QY      122 PheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGly 141
Db      664 GCGCGCGAATGGAGCGCGCGCCATGAGCGAGTCTGTCC----- 705
QY      142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAla 161
Db      706 ---ACGCCAGCGCTCGGTGGCGGAATT----- 729
QY      162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArg 181
Db      730 -----CTGATCGCGCAAG 741
QY      182 PheValProPheAlaAlaVal-----AlaAlaAlaAlaCysIle 194
Db      742 CTGCGCGCGGTACTTGTGCTACTCGGCATGCTGTCCACCGTGGGTGCGCGCGCGGTG 801
QY      195 AsnIleProLeuMetArgGlnArgGluLeuGlnValGlyIleProValAlaAspGluAla 214
Db      802 TTTCGTG-----TTTCGCGGTCCCGAATG----- 822
QY      215 GlyGlnArgLeuGlyTyrSerValThrAlaAlaLysGlnGlyIlePheGlnValValIle 234
Db      823 -----CGCGGCTCTGTGTCCTCGTGGCTC 849
QY      235 SerArgIleCysMetAlaIleProAlaMetAlaIleProProLeuIleMetAspThrLeu 254
Db      850 TTGTGCGCGCGTTCATGATGTCGCGCGCGGT---CAAGGCTGTGTAATTTCCTATG 906
QY      255 GluLysLysAspPheLeu 260
Db      907 GCGCGCAACCAAGTTCTCTC 924

```

```

RESULT 9
US-09-252-991A-2024/C
: Sequence 2024, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,768
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 2024
: LENGTH: 552
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2024

Alignment Scores:
Pred. No.:          0.117          Length:          552
Score:              85.00          Matches:         39
Percent Similarity: 34.76%        Conservative:    26
Best Local Similarity: 20.86%      Mismatches:     48
Query Match:        6.32%          Indels:         74
GB:                 4              Gaps:           8

```

```

US-09-990-415a-2 (1-261) x US-09-252-991A-2024 (1-552)
OY 35 ThrAspProArgAsnLeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleVal 54
   :||||| :||||| :|||||
Db 486 ACCGAGCCGCTGCTGACATGCTCACACCATGGCTCTACCCCAATACCCCACTACGTC 427
   :||||| :||||| :|||||
OY 55 GlnAsnTyrArgAlaGlyValValThrProGlyIleThrGluAspGlnLeuTyrArgAla 74
   :||||| :||||| :|||||
Db 426 GAGACTACGACCCGCGGCTGCTC-----CAGTTCTGGGCGCGC 388
   :||||| :||||| :|||||
OY 75 LysTyrValTyrAspSerAla----- 81
   :||||| :||||| :|||||
Db 387 GGG-----CTGGATGCTGCGCGCGCGCGCAAGCCGTCATCTGGAGCCGCTACTGG 334
   :||||| :||||| :|||||
OY 82 PheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnValPro 101
   :||||| :||||| :|||||
Db 333 TTCAACCCCGACCTGGAAAGCCGCGCGCACTGATT-----CCGCGCGCAATCGCC 283
   :||||| :||||| :|||||
OY 102 MetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrValVal 121
   :||||| :||||| :|||||
Db 282 ATCGTACATGACCATCATGCGCACATGCTG-----ACTGCATGGGCTG 238
   :||||| :||||| :|||||
OY 122 PheTyrGlnTyrPylAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSerGly 141
   :||||| :||||| :|||||
Db 237 GCGCGCGAATGGAGCGCGCGCACCATGGAGCAGCTGCTCC----- 196
   :||||| :||||| :|||||
OY 142 AspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAla 161
   :||||| :||||| :|||||
Db 195 ---ACGCCAGCCTCGGTGCGCGCAAT----- 172
   :||||| :||||| :|||||
OY 162 ValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGlyArg 181
   :||||| :||||| :|||||
Db 171 -----CTGATCGGCAAG 160
   :||||| :||||| :|||||
OY 162 PheValProPhe-----AlaAlaValAlaAlaAla 191
   :||||| :||||| :|||||
Db 159 CTGCTGCGCTACTTGTACTCGCATGCTGTCACGCTGGCGGCGCGCTGCGCGGTG 100
   :||||| :||||| :|||||
OY 192 AsnCysIleAsnIleProLeu 198
   :||||| :||||| :|||||
Db 99 TTCGTGTTGCGGCTGCCGATG 79

```

```

FILING DATE: 03-JUL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP91/00375
FILING DATE: 22-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Helwege, James W.
REGISTRATION NUMBER: 28,808
REFERENCE/DOCKET NUMBER: 514853
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-415-1500
TELEFAX: 703-415-1508
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2855 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 170..2482
US-07-923-976-5

Alignment Scores:
Pred. No.: 2.07 Length: 2855
Score: 84.50 Matches: 54
Percent Similarity: 38.86 Conservative: 28
Best Local Similarity: 25.59 Mismatches: 76
Query Match: 6.28 Indels: 54
Gaps: 11

US-09-990-415a-2 (1-261) x US-07-923-976-5 (1-2855)
OY 61 ValValThrProGlyIleThrGluAspGlnLeuTyrArgAlaLys-----TyrValTyr 78
   :||||| :||||| :|||||
Db 1676 ATCGTACTCC---TTGACCAAGACACCATGGAGCCCTCCAGCATGTCTATGCTTAC 1732
   :||||| :||||| :|||||
OY 79 AspSerAlaPheHisProAspThrGlyGlyLysValValLeu-----IleGlyArgMet 96
   :||||| :||||| :|||||
Db 1733 TCTCAAGAAATGGCTCCCTCCCATGCCCGCCAGAGCTGCATCTAAGACATTTGGCAAGAC 1792
   :||||| :||||| :|||||
OY 97 SerAlaGln-----ValProMetAsnMetThrIleThrGlyCysMetLeuThrPhe 113
   :||||| :||||| :|||||
Db 1793 TGGGACACACTGAGTGGCTGCTGAGCCCTGACCTGGGAGAGCCCTTCTCCGCC 1852
   :||||| :||||| :|||||
OY 114 TyrArgLysThrProThrValValPheTyr-----GlnTyrValAsnGlnSerPheAsnAla 132
   :||||| :||||| :|||||
Db 1853 TAC-----ACCATCTTGTGACCAACGCTCAGAACCATGCTTCTCCGCC 1897
   :||||| :||||| :|||||
OY 133 IleValAsnTyrSerAsnArgSer-----GlyAspThrProIleThrValArg 148
   :||||| :||||| :|||||
Db 1898 ATCTGAAATGCTCTCCCGCTGCTTGTCTTCACAGGCTGGAGCCCGCAGCTGTAT 1957
   :||||| :||||| :|||||
OY 149 GlnLeuGlyThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAlaLeuGlyLeu 168
   :||||| :||||| :|||||
Db 1958 CACATTCACCTATGCTGCTGCGCAGCGCTGGGGCCACCAACAGTACATCTCCACCTCG 2017
   :||||| :||||| :|||||
OY 169 LysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPheAlaAlaVal 188
   :||||| :||||| :|||||
Db 2018 ATGACCTTGACC-----CCAGCCCCACACAGGAAGA----- 2047
   :||||| :||||| :|||||
OY 189 AlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlnLeuGlnValGlyIle 208
   :||||| :||||| :|||||
Db 2048 -----ATCCCTCTGGCCA- 2061
   :||||| :||||| :|||||
OY 209 ProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerVal---ThrAlaAlaLysGln 227
   :||||| :||||| :|||||
Db 2062 ACTGTCCACAGCCACCATCAGACGCTGGCTCTCTGGCCCAATCATATGAGGAG 2121
   :||||| :||||| :|||||
OY 228 GlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAlaIlePro 247
   :||||| :||||| :|||||
Db 2122 GATGCTTCCAG-----CTGCGCGGCTTGGACACGCCA 2154
   :||||| :||||| :|||||

```

```

OY      248 ProteInlemerAsPThreundLUlystysasp 258
Db      2155 CCCATCACCACAGCTCACAGTGTCTGGAAGAGGATT 2187

RESULT 11
US-08-247-901C-1
Sequence 1, Application US/08247901C
Patent No. 5750384
GENERAL INFORMATION:
APPLICANT: Jacobs et al
TITLE OF INVENTION: L5 SHUTTLE PHASMIDS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rochstein & Edestein
STREET: 90 Park Avenue
City: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage disket
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/247,901C
FILING DATE: May 23, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/057,531
FILING DATE: April 29, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Elizabeth A
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 96700/273
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 50341
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: L5 shuttle plasmid sequence
HYPOTHETICAL: NO
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: L5 mycobacteriophage
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HALOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION: No. 5750384e
AUTHORS:
TITLE:
JOURNAL:
VOLUME:

```

DATE:	DOCUMENT NUMBER:	FILING DATE:	PUBLICATION DATE:	RELEVANT RESIDUES IN SEQ ID NO:
US-08-247-901C-1				
Alignment Scores:				
Pred. NO.:	512	Length:	50341	
Score:	82.00	Matches:	61	
Percent Similarity:	32.66%	Conservative:	36	
Best Local Similarity:	20.54%	Mismatches:	92	
Query Match:	6.10%	Indels:	108	
DB:	1	Gaps:	13	
US-09-990-415A-2 (1-261) x US-08-247-901C-1 (1-50341)				
QY	44	GLYALGILNLEUQILASERARGASNIIEVALGINSNTYRARGHLAGLYVAL	---	62
DB	16515	GGCCCGATGCTGACCCGGATCACCCCAACCTGCAGAGAGTGAGCTCGGAGCTCGGAC	16574	
QY	63	-----ThrProGlyLLeIhrGILuSPrGILNleu	71	
DB	16575	ATGAGCCGGGTGATCACCGAGCTGATCACCCAGGCTCGGTGTGAGAGATTCAGAAC	16534	
QY	72	TRPARGALALYSITGYVALTYRASPERRALPHEIAPROAPRTHIGLYLUSVAL	91	
DB	16635	ATGCTCACCAAG-----ACGGAGAGATTCTTCACG	16666	
QY	92	LEUILLEGIYARGMETSERALAGIIVALPROMETASMETHTHLETHRGLYSMETLEU	111	
DB	16665	GGGCTGGGCT-----CCTGTGCTGCTACCGGACGAGGCGTTCCTG	16708	
QY	112	ThrPhe-----	113	
DB	16707	ACGCTGTCCAAAGCCGGCGGAACCTGTCGGACGCTCTGGCTCCCTGCAGAGATTTC	16766	
QY	114	-----TYRARGLYSTHPROTHYVALPHE-----	122	
DB	16767	ACCAAGGGCTTCAACAGACATGTGCACACCGAGTCACGTCCACAGGCGCTTTCAGAGGTGCC	16826	
QY	123	TRPGILNTPYVALASNIINSERPHEASNIILIEVALASNTYSERASARG-----	139	
DB	16827	ATGCAAGGGCTTTCGGAGACGCTGGGACGCTCTCAACCTGTTCACCGGCTCATGTGAG	16886	
QY	140	-----SERGLYASPTHPRPROILETHRYVALARGILNleu	150	
DB	16887	TCCGGTCTGACGGCGATGGGACACGTCGGGCGGTCCGCTTCGACGTTCTCAACGGTTTC	16946	
QY	151	GLYTHRALATYRVALSER-----ALATHTHGLYVALVALAETHR	164	
DB	16947	GGAGATCTCTCTCTCTGCTGATGCCGGCGCTGACTCTCGTCTGCTGTGATTCGGCAAC	17006	
QY	165	ALALEUGILYLEUYSERLEUTHLYSHLSLEUPROPROLEUVALGILYARGPHEVALP	184	
DB	17007	GTCCTGGG-----ACGCTGGGACACACGCTGCTCCATGTGTACAGGCGCTCACGGCG	17060	
QY	185	---PHEALALAYALALALALASNCYSILIASNIILEPROLEUMETARGILNARGILU	203	
DB	17061	GGCTTCCAGACGCTGGCGAGC-----ACGCTGGGACAGATGCTCACGGAGGCC	17108	
QY	204	LEUGILNVALGILYLEUPROVALALASPRGILUALAGILNARGLEUGILYTYRSERVALTHR	223	
DB	17109	CTTCAAGCTCTGGGCTCGCATCTGACTACGTTGCTACGTTGATGCGCACGAGCTGAAC	17168	
QY	224	ALALALYSGLINLYLE-----PHEGIN-----	231	
DB	17169	ACGGCGCTGACGGCTTCCAGCCGATGCGCGCTGCTATGACAGAGCTTCCAGAGATTC	17228	
QY	232	-----VALVALLESERARGILECYSMETALALEPROIALMETALA-----	245	
DB	17229	TCCGAGCTATCTGGTGACAGCTGTGGCGCCGACATCCGCGCTGCGGACGGCCCTCGGC	17288	

Db 17229 TCCGACGTACTGTGTGACAGTCTGCCCCGACATCCCGGCGCTGGGAGAGCCCTCGGC 17288
 Oy 246 -----IlleProbleulellemetaphrleu 254
 Db 17289 CAGGTCCGAGCGCGGTGCTGACGTCTCGCAGCATCTCGACGTG 17339

RESULT 13
 US-09-426-436-1
 ; Sequence 1, Application US/09426436
 ; Patent No. 6225066
 ; GENERAL INFORMATION:
 ; APPLICANT: William R. Jacobs, Jr.
 ; APPLICANT: Barry R. Bloom
 ; APPLICANT: Graham F. Hatfull
 ; TITLE OF INVENTION: MYCOBACTERIAL SPECIES-SPECIFIC
 ; TITLE OF INVENTION: REPORTER MYCOBACTERIOPHAGES
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Amster, Rothstein & Edestein
 ; STREET: 90 Park Avenue
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10016
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch 1.44 MB storage diskette
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WORD PROCESSOR (ASCII)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/426,436
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/705,557
 ; FILING DATE:
 ; APPLICATION NUMBER: US/08/057,531
 ; FILING DATE:
 ; APPLICATION NUMBER: 07/833,431
 ; FILING DATE: February 7, 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pasqualini, Patricia A.
 ; REGISTRATION NUMBER: 34,894
 ; REFERENCE/DOCKET NUMBER: 96700/238
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 697-5995
 ; TELEFAX: (212) 286-0854 or 286-0082
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 52297
 ; TYPE: nucleotide
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE:
 ; DESCRIPTION: phage genome sequence
 ; HYPOTHEICAL: no
 ; ANTI-SENSE: no
 ; FRAGMENT TYPE: not applicable.
 ; ORIGINAL SOURCE:
 ; ORGANISM: mycobacteriophage L5
 ; STRAIN: not applicable
 ; INDIVIDUAL ISOLATE: L5
 ; DEVELOPMENTAL STAGE: not applicable
 ; HAPLOTYPE: not applicable
 ; TISSUE TYPE: not applicable
 ; CELL TYPE: not applicable
 ; ORGANELLE: not applicable
 ; IMMEDIATE SOURCE: mycobacteriophage L5 particles
 ; POSITION IN GENOME: entire genome
 ; FEATURE:
 ; NAME/KEY:

LOCATION:
 IDENTIFICATION METHOD:
 OTHER INFORMATION:
 PUBLICATION INFORMATION:
 AUTHORS: Hatfull and Sarkis
 TITLE: DNA sequence, structure and gene
 TITLE: Expression of Mycobacteriophage L5:
 TITLE: A Phage System for Mycobacterial
 TITLE: Genetics
 JOURNAL: Molecular Microbiology
 VOLUME: 7
 PAGES: 395-405
 DATE: 1993
 US-09-426-436-1

Alignment Scores:
 Pred. No.: 546
 Score: 82.00
 Percent Similarity: 32.668
 Best Local Similarity: 20.548
 Query Match: 6.10%
 Gaps: 13
 Length: 52297
 Matches: 61
 Conservative: 36
 Mismatches: 92
 Indels: 108

US-09-990-415A-2 (1-261) x US-09-426-436-1 (1-52297)

Oy 44 GLYAlaGlnLeuGluAlaSerArgAsnIleValAlaAsnTyrArgAlaGlyValAla--- 62
 Db 16412 GGCCCGATGCTGACCGGATGATCCACCCCACTGACGAGACGTGGCGCTGCTGTAAC 16471
 Oy 63 -----ThrpGlyIleThrGluAspGlnLeu 71
 Db 16472 ATGGCCGGGTCGATCACCGACGTGATACCCAGGCTCTGCTGACGACATCAAGAAC 16531
 Db 16532 ATCCTGACCAAG-----ACCGAGAGTTCTTACG 16561
 Oy 92 LeuIleGlyArgMetSerAlaGlnValProMetAsnMetThrIleThrGlyCysMetLeu 111
 Db 16562 GGCCCTCGGC-----CCTGTGCTGCTACCGGACGACGAGCGCTTCTG 16603
 Oy 112 ThrPhe----- 113
 Db 16604 ACCGCTGCCAAGCCGGCGGAACTGCTGGCAGCCTCTGCTCCCTGACAGAGATTC 16663
 Oy 114 -----TyrArgLysThrProThrValAlaPhe----- 122
 Db 16664 ACCAAGGCTTCAACGACATGTCACACCGAGTCAACGCGGTTCGAGGGTCC 16723
 Oy 123 TrpGlnTyrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArg----- 139
 Db 16724 ATGCAAGGGCTTTCGACAGCGTGGGAGCGTCTCAACCTGTTCACCGCGTCATGAG 16783
 Oy 140 -----SerGlyAspThrProIleThrValArgGlnLeu 150
 Db 16784 TCCGCTCGACGCGATGGAGACGCTCGCGGCTCCGCTGACGCTTCACACGAGGTTCC 16843
 Oy 151 GlyThrAlaTyrValSer-----AlaThrThrGlyAlaValAlaThr 164
 Db 16844 GGAATCTCTGCTGCTGCGATGCGGGCGCTGACTGCTGCTGCTGATGAGGCAAC 16903
 Oy 165 AlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValPro 184
 Db 16904 GTCCCTGGG-----ACGCTGGGACACAGACTGCTCCATCTCAACGCGCTCACGCCG 16957
 Oy 185 ---PheAlaAlaValAlaAlaAsnGlyIleAsnIleProLeuMetArgGlnArgGlu 203
 Db 16958 GCCTTCACACGCTGCGAGC-----ACCTGGGACGAGTCTCACCGGACCC 17005
 Oy 204 LeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyrSerValThr 223
 Db 17006 CTCCAAGCTCTGGTCCGATCTCGATGAGTGTGCTGCTAGCTGCTGCGACGACGCTGAC 17065

QY 224 AAlaAlaLysGlnGlyLeu-----PheGln----- 231
 Db 17066 ACAGCGCTGCTCCAGCCGATGCTGCGCTGCATGACAGACTTCAGAGATC 17125
 QY 232 -----ValValIleSerArgIleCysMetalalleProAlaMetala----- 245
 Db 17126 TCCGACGACTGCTGACAGCTGCTGCGCCGACATCCGCGCTGCGACGCGCCCTCGGC 17185
 QY 246 -----IleProProleuIleMetAspThrLeu 254
 Db 17186 CAGTCCGACGCGCTGCTGACGCTCGCTCCGACGATCATCTCGACGTTG 17236

RESULT 14
 ; Sequence 1, Application US/08705557
 ; Patent No. 6300061
 ; GENERAL INFORMATION:
 ; APPLICANT: William R. Jacobs, Jr.
 ; APPLICANT: Barry R. Bloom
 ; APPLICANT: Graham F. Hatfull
 ; TITLE OF INVENTION: MYCOBACTERIAL SPECIES-SPECIFIC
 ; TITLE OF INVENTION: REPORTER MYCOBACTERIOPHAGES
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Amster, Rothstein & Ebenstein
 ; STREET: 90 Park Avenue
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10016
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: Word Processor (ASCII)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/705,557
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/057,531
 ; FILING DATE:
 ; APPLICATION NUMBER: 07/833,431
 ; FILING DATE: February 7, 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pasqualini, Patricia A.
 ; REGISTRATION NUMBER: 34,894
 ; REFERENCE/DOCKET NUMBER: 96700/238
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 697-5995
 ; TELEFAX: (212) 286-0854 or 286-0082
 ; TELEX: TWX 710-581-4766
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 52297
 ; TYPE: nucleotide
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE:
 ; DESCRIPTION: phage genome sequence
 ; HYPOTHEICAL: no
 ; ANTI-SENSE: no
 ; FRAGMENT TYPE: not applicable.
 ; ORIGINAL SOURCE:
 ; ORGANISM: mycobacteriophage L5
 ; STRAIN: not applicable
 ; INDIVIDUAL ISOLATE: L5
 ; DEVELOPMENTAL STAGE: not applicable
 ; HAPLOTYPE: not applicable
 ; TISSUE TYPE: not applicable
 ; CELL TYPE: not applicable
 ; CELL LINE: not applicable
 ; ORGANELLE: not applicable

IMMEDIATE SOURCE: mycobacteriophage L5 particles
 POSITION IN GENOME: entire genome
 FEATURE:
 NAME/KEY:
 LOCATION:
 IDENTIFICATION METHOD:
 OTHER INFORMATION:
 PUBLICATION INFORMATION:
 AUTHORS: Hatfull and Sarkis
 TITLE: DNA Sequence, Structure and Gene
 TITLE: Expression of Mycobacteriophage L5:
 TITLE: A Phage System for Mycobacterial
 TITLE: Genetics
 JOURNAL: Molecular Microbiology
 VOLUME: 7
 PAGES: 395-405
 DATE: 1993
 US-08-705-557-1

Alignment Scores:
 Pred. No.: 546 Length: 52297
 Score: 82.00 Matches: 61
 Percent Similarity: 32.66% Conservative: 36
 Best Local Similarity: 20.54% Mismatches: 92
 Query Match: 6.10% Indels: 108
 DB: 4 Gaps: 13

US-09-990-415a-2 (1-261) x US-08-705-557-1 (1-52297)

QY 44 GlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyArgAlaGlyValaL--- 62
 Db 16412 GGCCGATGCTGACCGGATCATCCCAACCTGACGAGAGCGGCTCGTGGTCAAC 16471
 QY 63 -----ThrProGlyIleThrGlnAspIleLeu 71
 Db 16472 ATGGCCGGGTGATCACCAGCTGATCACCAGGCTCTGCTGTCACAGATCCAGAAC 16531
 QY 72 TrpArgAlaLysTyValTyAspSerAlaPheHisProAspThrGlyIleValaVal 91
 Db 16532 ATCCACACCAAG-----ACCGAGAGTCTTCACG 16561
 QY 92 LeuIleGlyArgMetSerAlaGlnValProMetAsnMetThrIleThrGlyCysMetLeu 111
 Db 16562 GGCGTCGGC-----CCTGTGCTCGCTACCGGACGACGACGCGCTTCCTG 16603
 QY 112 Thrphe----- 113
 Db 16604 ACCTGTCCAAAGCCCGCGCACTGTTGCGCACGCTCTGGCTCCCTCGACGAGTTTC 16663
 QY 114 -----TyrArgLysThrProThrValaValPhe----- 122
 Db 16664 ACCAAGCGCTTCACAGCATGTGTCACAGATGTCACAGCGCTTCGAGAGGCTGCC 16723
 QY 123 TrpGlnTrpValaAsnGlnSerPheAsnAlaIleValaAsnTySerAsnArg----- 139
 Db 16724 ATGCAAGGCTTTCGACAGACCTGGCGCCTCTCCACAGCTTCACAGCGCTCATGAG 16783
 QY 140 -----SerGlyAspThrProIleThrValaArgGlnLeu 150
 Db 16784 TCCGCTGACAGCGCATGGACAGCTCGGCGCTCGCTGACGCTTCATCAACGCGTTTC 16843
 QY 151 GlyThrAlaTyValSer-----AlaThrThrGlyAlaValaAlaThr 164
 Db 16844 GGAGATCTCTGCTGCTGCTGATGCGCGGCTGATCTGCGTCTGGTGTATCGCAAC 16903
 QY 165 AlaIleuGlyLeuLysSerLeuThrLysIleuProProleuValaGlyArgPheValPro 184
 Db 16904 GTCTCTCGGG-----ACGCTGCGCACACAGCTCGCTCCATCATGTCACGCGCTCAC 16957
 QY 185 ---PheAlaAlaValaAlaAlaAsnGlyIleAsnIleProleuMetArgGlnArgGlu 203
 Db 16958 GCCTTCAGACGCTGGCGAGC-----ACGCTCGGACGATGCTCACCGGAGCC 17005

QY 204 LeuGlnValGlyIleProValAlaAspGluAlaGlyInArgLeuGlyTyrSerValThr 223
||||| |||||
DB 17006 CTCGAAGCTCTGGGTGGATCCTGACTGACGTGCGTACTGTGATGCGCACGACCTGAAC 17065
QY 224 AlaAlaLysGlnGlyIle-----PheGln----- 231
DB 17066 ACGGCGCTGTCAGGCTCTCCAGCGATGCTGCCCTGCTCATGACAGATTCACAGAGATC 17125
QY 232 -----ValValIleSerArgIleCysMetAlaIleProAlaMetAla----- 245
DB 17126 TCCGACGACTGCTGACGACGCTGGCCCCGACATCCGCGGCTGGCGACGCGCTCGGC 17185
QY 246 -----IleProLeuIleMetAspThrLeu 254
DB 17186 CAGTGCAGCGGCGGTGCTGACGCTCGCTCCGACGATCATCTCGACGTTG 17236
RESULT 15
US-09-252-991A-9294/c
; Sequence 9294, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9294
; LENGTH: 663
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9294
Alignment Scores:
Pred. No.: 0.526 Length: 663
Score: 81.00 Matches: 35
Percent Similarity: 39.60% Conservative: 24
Best Local Similarity: 23.49% Mismatches: 52
Query Match: 6.02% Indels: 38
DB: 4 Gaps: 6
US-09-990-415a-2 (1-261) x US-09-252-991A-9294 (1-663)
QY 118 ProThrValValPheTrpGlnTrpValAsnGlnSerPheAsnAla-----Ile 133
||| |||||
DB 484 CCGGGCGTGTGGCTGTGCTGGCGCTGCTGCGCTGCTGCGCTGCGCTGCGCTGCTGCTG 425
QY 134 ValAsnTyrSerAsnArgSerGlyAspThr-----ProIleThrValArgGlnLeuGly 151
||| |||||
DB 424 GCGGCTTACGCTACGCGCGCGGATCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 365
QY 152 ThrAlaTyrValSerAlaThrThrGlyAlaValAlaThrAlaLeuGlyLeuLysSerLeu 171
: ||| : |||
DB 364 AGCGGCTGAACGCGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCGCTG 308
QY 172 ThrLysIleuProProLeuValGlyArgPheValProPheAlaIleValAlaIleAla 191
: ||| : |||
DB 307 -----ATGCCCTCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 281
QY 192 AsnCysIleAsnIleProLeuMetArgGlnArgGlnLeuGlnValGlyIleProValAla 211
||| |||||
DB 280 CTCGTCACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 221
QY 212 AspGluAlaGlyInArgLeuGlyTyrSerValThrAlaAlaLysGlnGlyIlePheGln 231
||| |||||
DB 220 -----CGGCTG-----TGGCTGCTGAAGCGCTGCGCATGCGACTGCGCGCC 179
QY 232 ValValIleSerArgIle----- 237

DB 178 GGCCTGCTCTTCCGATTCGACGCGCATCTGCTGCAACTGGGCTGGGCACTGCGCGGC 119
QY 238 CysMetAlaIleProAlaMetAlaIle 246
DB 118 TGCCTGCGGCTGCTGCTGCGCATGCGCGCTC 92
Search completed: October 16, 2003, 17:36:46
Job time: 132 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - nucleic search, using frame_plus.p2n model

Run on: October 16, 2003, 14:00:26 ; Search time 263 Seconds

(without alignments)
2678.910 Million cell updates/sec

Title: US-09-990-415A-2

Perfect score: 1345

Sequence: 1 MESKMGELPLDINIOEPRMD.....PAMAIPLIMDTLEKDFLK 261

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-O=/cgn2.1/USPTO.spool_p/US09999415/runat.16102003.073806.9430/app_query.fasta.1.455
-DB=N.Geneseq.19jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LIST=45 -DOCALLIGN=200 -THR.SCORE=PCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09990415.BCGN.1.1.396 @runat.16102003.073806.9430 -NCPU=6 -ICPU=3
-NO.MMAP -LARGESOURCE -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGIO
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :
N.Geneseq.19jun03:*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1345	100.0	1232	24	ABN83754	Human protein clus
2	1345	100.0	1375	24	ABO98865	Human ORF672 codin
3	1345	100.0	1729	22	AAI59123	Human polynucleoti
4	1345	100.0	1826	22	AAE27703	Human transport pr
5	1345	100.0	2702	22	AAK94713	Human full-length
6	1345	100.0	3047	22	AAH72722	Human cervical can
7	1318	98.0	984	24	ABL39805	Human NS cDNA sequ
8	1269	94.3	843	21	AAK75798	Human ORFX ORF1353
9	1223.5	91.0	2215	23	AA584362	DNA encoding novel
10	1089	81.0	1076	22	AAI59569	Human polynucleoti
11	1089	81.0	1134	24	ABN84365	Alzheimer's diseas
12	1089	81.0	1485	22	AAI26555	Human breast canc
13	1089	81.0	2053	22	AAI61354	Human polynucleoti
14	1089	81.0	2269	24	ABN83757	Human protein clus
15	1089	81.0	2367	24	AAD21995	Human transporters
16	1089	81.0	2468	21	AAK75794	Human ORFX ORF1349
17	1082	80.4	2947	22	AAH44832	Rat cDNA encoding
18	1053	78.3	752	22	AAK91763	Human cDNA 5'-end
19	1053	78.3	752	22	AAK93199	Human cDNA clone r
20	1051.5	78.2	1503	24	ABK54156	cDNA encoding huma
21	1026	76.3	2766	22	AAH15610	Human cDNA sequenc
22	1003.5	74.6	1061	24	ABN83755	Human protein clus
23	961.5	71.5	1084	24	ABK96715	DNA encoding human
24	897.5	66.7	984	24	ABK96716	DNA encoding splic
25	897.5	66.7	984	24	ABN84366	Alzheimer's diseas
26	798	59.3	2642	22	AAH17727	Human cDNA sequenc
27	765	56.9	741	22	AAH06089	Human cDNA clone (
28	754	56.1	1286	23	ABL18135	Drosophila melanog
29	754	56.1	1567	24	ABN83756	Human protein clus
30	750	55.8	1772	24	AAD33666	Human TRICH-21 cDN
31	750	55.8	2312	22	AAI59647	Human polynucleoti
32	750	55.8	2376	22	AAI57861	Human polynucleoti
33	750	55.8	2458	24	ABK35352	Human cDNA encodin
34	712.5	53.0	1575	24	ABL98804	Human secretory po
35	693	51.5	806	22	AAH04784	Human cDNA clone (
36	670	49.8	507	23	ABV45979	Human prostate exp
37	626	46.5	421	23	ABV16181	Human prostate exp
38	602	44.8	524	22	AA531167	Human diagnostic a
39	599.5	44.6	1585	23	ABL11787	Drosophila melanog
40	578	43.0	4258	23	ABL18134	Drosophila melanog
41	563	41.9	4438	23	ABL1786	Drosophila melanog
42	545	40.5	413	23	AA584361	DNA encoding novel
43	480.5	35.7	2673	24	AAD36304	Human transporter
44	467	34.7	1244	20	AAK97798	Extended human sec
45	462	34.3	1548	24	ABL90501	Human polynucleoti

ALIGNMENTS

RESULT 1
ABN83754
ID ABN83754 standard; cDNA; 1232 BP.

AC ABN83754;

DT 02-SEP-2002 (first entry)

DE Human protein cluster I gene.

KM Protein cluster I; human; metabolic disorder; obesity; diabetes;

OS antidiabetic; diagnosis; therapy; chromosome 10; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 450..1232

```

FT      /*tag= a /product= "Protein cluster I"
FT      /partial
FT      /note= "the CDS does not include a stop codon"
PN      WO200242324-A1.
PD      30-MAY-2002.
XX
XX      22-NOV-2001; 2001WO-SE02581.
PR      24-NOV-2000; 2000SE-0004325.
XX
XX      (PHAA ) PHARMACIA AB.
PI      Attersand A;
XX
XX      WPI; 2002-500277/53.
DR      P-PSDB; ABB76445.
PT
PT      Novel nucleic acid molecule encoding Protein Cluster I, useful in the
PT      diagnosis of metabolic diseases, such as obesity and diabetes, and in
PT      the identification of agents useful in the treatment of the diseases
PS      Claim 1; Page 23-24; 34pp; English.
XX
XX      The present invention relates to the identification of a human gene
CC      family (see ABN83754-57) encoding a group of polypeptides (see
CC      ABB76445-48) referred to as Protein Cluster I. This family of
CC      homologous proteins was identified by an "all-versus-all" BLAST
CC      procedure using all Caenorhabditis elegans proteins in the
CC      Wormpep20 database release. Proteins of unknown function were
CC      compared to the Drosophila melanogaster Flybase database, and
CC      non-annotated protein clusters, conserved in both C. elegans and D.
CC      melanogaster, were used in a BLAST procedure against the Celera
CC      Human Genome Database, and Protein Cluster I proteins of unknown
CC      function were selected for study. The human part of Protein
CC      Cluster I comprises polypeptides encoded by 3 genes; an alternative
CC      splicing of the present sequence (deletion of nucleotides 624-794)
CC      gives the sequence in ABN83755. The present gene was identified in
CC      a human DNA sequence from clone RP11-1081.7 on chromosome 10. It
CC      is expressed primarily in the nervous system and digestive system.
CC      The claimed Protein Cluster I nucleic acid molecules and proteins
CC      are proposed to be useful for differential identification of the
CC      tissue(s) or cell type(s) present in a biological sample, for
CC      diagnosis of diseases and disorders, including metabolic disorders
CC      and immune disorders, especially obesity and diabetes, and for
CC      identifying agents useful in the treatment of such diseases. The
CC      nucleic acid molecules are also useful as hybridisation probes, for
CC      chromosome and gene mapping, in PCR technologies, and in the
CC      production of sense or antisense nucleic acids.
XX
XX      Sequence 1232 BP; 239 A; 368 C; 375 G; 250 T; 0 other;
SQ
SQ      Sequence 1232 BP; 239 A; 368 C; 375 G; 250 T; 0 other;

Alignment Scores:
Pred. No.:      8. 01e-143      Length:      1232
Score:          1345.00      Matches:      261
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             Gaps:        0

US-09-990-415A-2 (1-261) x ABN83754 (1-1232)
OY      1 MetGUserIyMeGlyGluUeuProUeUsApIleAsnIleGInGluPProAlyTTPAsp 20
Db      450 ATGGAAAGCAAAATGGGTCAATTTGGCTTTAGACATCAACATCCAGAAACCTCGTGGAC 509
OY      21 GInSerThPhLeUeGlyYArGAlaIArgHsPhePheThValIThrAsPProArgSnLeu 40
Db      510 CAAGAGACTTCTCGGCGAGACC CGCGACACTTTTCACTGTACTGATCCTCGAAATCTG 568
OY      41 LeuUeuSerGlyAlaGInLeUeGInUaIaSerArGAsnIleValGInAsnTyTArGAlaGly 60

```

Dd		570	CTGCGTGTGGGGGCACAGCTGGAAGCTTCCTCGGAACATCGTGACAACTACAGGGCCGGC	629
Oy		61	VAlValThrProGlyIleThrGluAspIleLeuTrpArgAlaLysTyValItyrAspSer	80
Dd		630	GTTGTGACCCCAGGAGATCACCGAGACCAGCTGTGGAGGGCCAATATGTTATGATCAACC	689
Oy		81	AlaPheHisProAspThrGlyGlyLysValValIleuIleGlyArgMetSerAlaGlnVal	100
Dd		690	GCCTTCCTCATCCGGACACAGGGGAGAGAAGTGGTCTCCGATTGGCCGATGTACGCCAGGTG	749
Oy		101	PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyArgLysThrProThrVal	120
Dd		750	CCCCATGAACATGACCATCATCTACTGGCTGCATGCTCACTTACATCAGAGAAACCCCAACCGTGT	809
Oy		121	ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer	140
Dd		810	GTTGTTCTGGCAGTGGGTGAATCAGTCTTCAATGCCATGTTAATCTCAACACGCGAGT	869
Oy		141	GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrGly	160
Dd		870	GGTGACACTCCCATCATCAGTAGGACAGCTGGGGACAGCTATGTAGTGCACCACTGGA	929
Oy		161	AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly	180
Dd		930	GCTGTGGCCAGCGCCCTGGGACTCAAATCCCTCACCAAGACACTGGCCCCCTTGGTGGCG	989
Oy		181	ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg	200
Dd		990	AGATTGTGGCCCTTTGGCAGCGATGGGACGCTCCAACTGCATCAACATCCCTGATGAG	1049
Oy		201	GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr	220
Dd		1050	CAGAGAGAGCGCAGGTGGGATCCCGGTGGCTGTGTGGCAGGACAAGGCTTGGCTTAC	1109
Oy		221	SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla	240
Dd		1110	TGCGGTGACTGACGACCAAGCAGGAAATCTCCAGGTGGTGATTTCAGAAATGTGCATGGCG	1169
Oy		241	IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysAspPheLeu	260
Dd		1170	ATTCTGTCGATGAGCCATCCACCACTGATGTAGGACACTGTGAGAAAGAAAGACTTCTCTG	1229
Oy		261	Lys 261	
Dd		1230	AAG 1232	
RESULT 2				
ABQ98865	ID	ABQ98865	standard; DNA; 1375 BP.	
XX AC	ABQ98865:			
XX DT	04-NOV-2002 (first entry)			
XX DE	Human ORF672 coding sequence.			
XX KW	Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;			
KW KM	Antiinflammatory; gene therapy; human; ORF; atherogenic; platelet;			
KW RW	human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;			
KM KW	cancer; cardiovascular disease; allergy; autoimmune disease;			
KW KM	wound healing; blood coagulation disorder; Inflammatory disorder; ds.			
XX OS	Homo sapiens.			
XX PN	US2002082206-A1.			
XX PD	27-JUN-2002.			
XX PF	30-MAY-2001; 2001US-0867550.			
XX PR	30-MAY-2000; 2000US-208427P.			

PA (LEAC/) LEACH M D.
 PA (MEHR/) MEHRABAN F.
 PA (CONL/) CONLEY P B.
 PA (TOPP/) TOPPER J N.
 PA (LAWD/) LAW D.
 PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
 XX WPI: 2002-626554/67.
 DR P-PSDB: ABP64302.
 XX
 PT New polypeptide designated ORFX are present in human atherogenic cells
 PT and are useful to prevent and treat ORFX-associated disorders including
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or
 PT inflammatory disease
 XX
 PS Claim 2; SEQ ID 1343; 78bp; English.
 XX
 CC The present invention relates to novel human ORFX polypeptides and their
 CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
 CC were discovered in human atherogenic cells, in particular in platelets
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
 CC many other tissues as well. Atherogenic cells are cells which have the
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
 CC nucleic acids are useful for treating or preventing a pathological
 CC condition associated with an ORFX-associated disorder, e.g. cancer,
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
 CC coagulation disorders or inflammatory disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPRO web site at seqdata.uspto.gov/sequence.html?DocID=20020082206.
 XX
 S0 Sequence 1375 BP; 342 A; 371 C; 374 G; 286 T; 2 other:
 Alignment Scores:
 Pred. No.: 9-4e-143 Length: 1375
 Score: 1345.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-990-415a-2 (1-261) x ABQ98865 (1-1375)
 QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrrPasp 20
 DB 91 ATGGAAGCAAAATGGGTGATGGCTTAGCATCAACATCCAGAACTCGCGGAGAC 150
 QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
 DB 151 CAAAGTACTTCTCTGGGAGAGCCCGGCACTTTTCACTGTACTGATCTCGAAATCTG 210
 QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 DB 211 CTGCTGTCGGGGGACACAGCTGGAAGCTTCTCGAACAATCGTGCAGAACTACAGGCCGCGC 270
 QY 61 ValValThrProGlyIleThrGluAspGlnLeuTrrPargAlaLysTyrValTyrAspSer 80
 DB 271 GTGTGACCCCAAGGATCAGCAGGACCAAGCTGTGGAGGCGCAAGTATGTATGATACACC 330
 QY 81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
 DB 331 GCCCTTCATCCGGACACAGGGGAGAGGTGTCTCGATGTGCGCGATCTCAGCCAGCTG 390
 QY 101 ProMetAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 DB 391 CCCATGAACATGACCATCTGCGCTGCATCTCATCTTCTACAGGAACACCCCAACCTG 450
 QY 121 ValPheTrrPglIntrrPvalAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 DB 451 GTGTTTGGCAGTGGGTGAATCAGTCTTCAATGATGCAATGTTAACTACTCCAAACCGCAGT 510
 QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160

DB 511 GGTGACACTCCCATCACTGTGAGCAGCGGGGAGACCTATGATGTGCCACCACTGGA 570
 QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
 DB 571 GCTGTGGCCACGCGCCCTGGGACTCAATCTCTCACCAACAGCACTGCCCCCTTGGTGGC 630
 QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnLysIleAsnIleProLeuMetArg 200
 DB 631 AGATTGTGCTCCCTTGGACAGCACTGGCAGCTGCCAATGCACTGCATCAATCCCTGATGAGG 690
 QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
 DB 691 CAGAGAGAGCTGACAGTGGGCACTCCCGGTGCTGATGAGCAGCTCAGAGGCTTGCTAC 750
 QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
 DB 751 TCGGTGACTGCAGCCAGCAGGGAATCTTCCAGTGGTGAATTCAGAAATCTGCATGGCG 810
 QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
 DB 811 ATTCCTGCATGGCCATCCACCACTGATCATGACACACTCTGGAGAAAGAAACTTCTCG 870
 QY 261 Lys 261
 DB 871 AAG 873
 RESULT 3
 AA159123
 ID AA159123 standard; cDNA; 1729 BP.
 XX AA159123;
 AC
 DF 22-OCT-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 1326.
 XX
 DE Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW Leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0596042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao Qa, Zhou P, Goodrich R, Dzmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR P-PSDB: AAM39967.
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -

PS Claim 1: SEQ ID NO 1326: 10078bp; English.
xx
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AA157798-AA161369) with neurotrophic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Actin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
xx
SQ Sequence: 1729 BP; 406 A; 469 C; 491 G; 363 T; 0 other;

Alignment Scores:
Pred. No.: 1,31e-142 Length: 1729
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-990-415A-2 (1-261) x AA159123 (1-1729)
QY 1 MetGluSerLysMetGluLysLeuProLeuAspLleAsnIleGlnGluProArgTyrPasp 20
DB 247 ATGGAAAGGAAATGGGTAATGCTTATGACATCAATCAAGAACCTCGCTGGAGC 306
QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPheThrValThrAspProArgAsnLeu 40
DB 307 CAAAGTACTTCTCTGGGAGAGCCCGCACTTTTACGTCATCTATCTCTCGAATCTG 366
QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
DB 367 CTGCTGTCGGGGGACAGCTGGAAGCTTCTCGAATCATGTCAGAACCTACAGGCGCGGC 426
QY 61 ValValThrProGlyLysThrGluAspGlnLeuThrArgAlaLysTyrValTyrAspSer 80
DB 427 GTGGTGACCCAGGAGATCACCGAGGACAGCTGTGGAGGCCAAGATATGTATGACTCC 486
QY 81 AlaPheHisProAspThrGlyLysValValLeuIleGlyArgMetSerAlaGlnVal 100
DB 487 GCCTTCATCCGAGACAGAGGAGAGAGTGTCTGATTTGGCCGATGTCCAGCCAGGTG 546
QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
DB 547 CCCATGACATGACATCTCATCTGCTGCTCATCTCATCTCATGAGGAGAGCCCAACCGTG 606
QY 121 ValPheTrpGlnIleThrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
DB 607 GTGTTCTGGCAGTGGGTGATCATGCTTCAATGCGCATTTTACTACTCAACCGCACT 666
QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
DB 667 GGTGACATCCCATCATCTGTGAGGAGCTGGGAGAGCGCTATGTAGTGTCCACCATGGA 726
QY 161 AlaValAlaThrAlaLeuGlyLysSerLeuThrLysHisLeuProProLeuValGly 180
DB 727 GGTGTGGCCAGGAGCCCTGGGACTCAATCTCCACCAACACCTGCCCCCTGGTGGCGC 786
QY 181 ArgPheValProPheAlaAlaValAlaAlaAsnCysIleAsnIleProLeuMetArg 200
DB 787 AATTTTGTGCTCTTTCACAGCAGTGGAGCTGCCAATGTCATCAATCTCCCTGATGAGG 846
QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220

DB 847 CAGAGAGACTGCGAGGTGGGCAATCCCGGTGCTGATGAGAGCAGCTCAGAGCTTGGCTAC 906
QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
DB 907 TCGGTGACATGAGCCCAAGAGGAGAACTTCCAGGTGTGTATTCAGAAATCTCATGCGC 966
QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeu 260
DB 967 ATTCTGCAATGACATCCACCATGATCATGACACTCTGGAGAAAGAACTTCTG 1026
QY 261 Lys 261
DB 1027 AAG 1029

RESULT 4
AAE27703
ID AAE27703 standard; cDNA; 1826 BP.
XX
AC AAE27703;
XX
DT 28-MAR-2001 (first entry)
XX
DE Human transport protein TPPT-3 coding sequence.
XX
KW Human: transport protein; TPPT; transport disorder; metabolic disorder;
KW neurological disorder; cardiovascular disorder; reproductive disorder;
KW immune disorder; cancer; ss.
XX
OS Homo sapiens.
XX
PN MO200078953-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000MO-US16668.
XX
PR 17-JUN-1999; 99US-0139923.
PR 10-AUG-1999; 99US-0148177.
PR 18-AUG-1999; 99US-0149357.
PR 28-OCT-1999; 99US-0162287.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzai Y, Lu DM, Au-Young J, Patterson C;
XX
DR WPI: 2001-041424/05.
DR P-PSDB; AAB60083.
XX
PT Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -
XX
PS Claim 5; Page 142-143; 165pp; English.
XX
CC The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TPPTs). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
XX
SQ Sequence 1826 BP; 415 A; 519 C; 512 G; 380 T; 0 other;

Alignment Scores:
Pred. No.: 1,42e-142 Length: 1826
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-990-415A-2 (1-261) x AAE27703 (1-1826)

QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
 DB 449 ATGGAAAGCAAAATGGGTGAATTCCTTTAGACATCAACATCCAGGAACCTCGCTGGGAC 508
 QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValIThrAspProArgAsnLeu 40
 DB 509 CAAAGTACTTCTCGGGGAGAGCCCGGACATTTTCACTGTACTGATGATCCCTCGAAATCTG 568
 QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 DB 569 CTGCTGTCCGGGACACAGCTGGAAGCTTCTCGGAACATCGTCGGAACCTACAGGGCCGCG 628
 QY 61 ValValThrProGlyIleThrGluAspGlnLeuThrPargAlaLysTyrValTyrAspSer 80
 DB 629 GTGGTGAACCCAGGAGATCACCGAGACAGCTGTGGAGGGCCAGTATGTATGACATCC 688
 QY 81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
 DB 689 GCTTCATCCGGACAGAGGGAGAGAGTGTCTGTGATGGCCCATGTCCAGCCAGGTG 748
 QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 DB 749 CCCATGAACATGACATCAGCTGCTGATCTCATCTTCACAGGAAGACCCCAACCTG 808
 QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 DB 809 GTGTTCTGGCGAGTGGTGAATCAGTCTCAATGCCATTTGTACTACTCCAACCCGACGT 868
 QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrIleValValSerAlaThrThyGly 160
 DB 869 GGTGACACTCCCATCTGTGAGGACAGCTGAGGACAGCTATGTGAGTGCACACACAGTGA 928
 QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
 DB 929 GCTGTGGCCAGGGCCCTGGGACTCAAAATCCCTCACAAGACCTGCCCCCTTGTGGTGGC 988
 QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnLysIleAsnIleProLeuMetArg 200
 DB 989 AGATTGTGCTGCTTGGAGCGAGTGGCAGCTGCCAAGCTGCATCAACATCCCTCATGAGG 1048
 QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluIleGlnArgLeuGlyTyr 220
 DB 1049 CAGAGAGAGCTGCGAGTGGGATCCCGTGGCTGATGAGGAGTCAAGGCTTGGCTAC 1108
 QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
 DB 1109 TCGGTGACTGAGCCAGCAAGGAGGAATCTTCCAGGTGGGATTTCAAGATGTGATGGCG 1168
 QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
 DB 1169 ATTGCTGCGATGGCCATCCCACTGATCATGTGACACTGTGAGAAAGAAAGACTTCTCTG 1228
 QY 261 Lys 261
 DB 1229 AAG 1231
 DB 1229 AAG 1231
 RESULT 5
 AAK94713
 ID AAK94713 standard; cDNA; 2702 BP.
 AC AAK94713;
 XX 06-NOV-2001 (first entry)
 DE Human full-length cDNA, seq ID NO: 3751.
 XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 OS Homo sapiens.
 XX
 OS Homo sapiens.
 XX
 PN EPI130094-A2.
 XX
 PD 05-SEP-2001.

XX 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX
 DR WPI; 2001-524255/58.
 DR P-PSDB; AAM93760.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 PS Claim 8; SEQ ID NO 3751; 1380bp + sequence listing; English.
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a full length
 CC human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX
 SQ Sequence 2702 BP; 608 A; 757 C; 699 G; 638 T; 0 other:
 Alignment Scores:
 Pred. No.: 2,51e-142 Length: 2702
 Score: 1345.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-990-415a-2 (1-261) x AAK94713 (1-2702)
 QY 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
 DB 70 ATGGAAAGCAAAATGGGTGAATTCCTTTAGACATCAACATCCAGGAACCTCGCTGGGAC 129
 QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValIThrAspProArgAsnLeu 40
 DB 130 CAAAGTACTTCTCGGGGAGAGCCCGGACATTTTCACTGTACTGATGATCCCTCGAAATCTG 189
 QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 DB 190 CTGCTGTCCGGGACACAGCTGGAAGCTTCTCGGAACATCGTCGGAACCTACAGGGCCGCG 249
 QY 61 ValValThrProGlyIleThrGluAspGlnLeuThrPargAlaLysTyrValTyrAspSer 80
 DB 250 GTGGTGAACCCAGGAGATCACCGAGACAGCTGTGGAAGGCCAAGTATGTATGACATCC 309
 QY 81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
 DB 310 GCTTCATCCGGGACAGAGGGAGAGTGTCTGTGATGGCCGATGTCAAGCCAGGTG 369
 QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 DB 370 CCCATGAACATGACATCAGCTGCTGATCTCATCTTCACAGGAAGAACCCCAACCTG 429
 QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 DB 430 GTGTTCTGGCGAGTGGTGAATCAGTCTTCAATGCCATTTGTACTACTCCAACCCGACGT 489

QY 141 GYASPTPrProIleThrValArgInLeuGlyThrAlaTyrValSerAlaThrThrGly 160
 DB 490 GGTGACACCTCCATACCTGTGAGCGACGCTGGGACAGCCTATGTGATGCCACCTGGA 549
 QY 161 AlaValAlaThrAlaLeuGlyLeuIleuYsSerIleuThrLysHisLeuProProLeuValGly 180
 DB 550 GCTGTGGCCAGCGCCCTGGGACTCAATCCCTCACCAAGACACTGGCCCCCTGGTGGCG 609
 QY 181 ArgPheValProPheAlaAlaValAlaAlaAsnCysIleAsnIleProLeuMetArg 200
 DB 610 AATTTTGTCTCCCTTTGACAGAGTGGGAGCTGCCACTGCATCAACATCCCTGTAGAGG 669
 QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
 DB 670 CAGAGAGAGCTCAGAGTGGGCAATCCAGTGGTGAAGAGCAGCTGAGGCTGGCTAC 729
 QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValAlaIleSerArgIleCysMetAla 240
 DB 730 TCGGTGACTGCAGCAAGCAGGGAATCTCCAGTGTGATTTCAAGAAATCTGACATGGCG 789
 QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
 DB 790 ATTCTGCGATGGCCATCCACCACTGATCATGACACTGTGAGAAAGAACTCTCG 849
 QY 261 Lys 261
 DB 850 AAG 852
 RESULT 6
 AAH72722/c
 ID AAH72722 standard; cDNA; 3047 BP.
 AC AAH72722;
 XX
 DT 19-SEP-2001 (first entry)
 XX Human cervical cancer marker nucleic acid 3996.
 DE
 XX
 KM Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 OS Homo sapiens.
 FN W0200142467-A2.
 PD 14-JUN-2001.
 XX
 PF 08-DEC-2000; 2000MO-US33312.
 XX
 PR 08-DEC-1999; 990US-0169681.
 PR 21-DEC-1999; 990US-0171350.
 PR 14-MAR-2000; 2000US-0189315.
 PR 12-MAY-2000; 2000US-0203791.
 PR 09-JUN-2000; 2000US-0210600.
 PR 21-JUL-2000; 2000US-0220114.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Deeds J, Berger A, Zhao X.
 XX
 DR WPI: 2001-375006/39.
 XX
 PT New isolated nucleic acid for diagnosing and treating cervical cancer
 CC and for assessing and detecting compounds for treating the cancer -
 CC
 PS Claim 1; Page 815-816; 1051pp; English.
 XX
 CC The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful: to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;
 CC and to select and/or assess the efficacy of a compound or therapy for
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be

CC useful for gene therapy.
 XX
 SQ Sequence 3047 BP; 629 A; 790 C; 891 G; 733 T; 4 other;
 Alignment Scores:
 Pred. No.: 2,99e-142 Length: 3047
 Score: 1345.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-990-415a-2 (1-261) x AAH72722 (1-3047)
 QY 1 MetGluSerLysMetGlyLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
 DB 2605 ATGGAAGCAAAATGGGTAAATGGCTTATGACATCAATCAGAGACCTGCTGGAC 2546
 QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValIleThrAspProArgAsnLeu 40
 DB 2545 CAAAGTACTTCTCTGGGAGAGCCCGGCACTTTTCACTGTACTGATCTCGAAATCTG 2486
 QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 DB 2485 CTGCTGTCCGGGGCAGACGCTGGAAGCTTCTCGGAACATGCTGCAGAACTACAGGGCGCG 2426
 QY 61 ValValThrProGlyIleThrGluAspGlnLeuTyrPargAlaLysTyrValIleThrAspSer 80
 DB 2425 GTGGTGACCCCGAGGATCACCGAGCCAGCTGTGGAGGCCAAGATGTGTATGACTCC 2366
 QY 81 AlaPheHisProAspThrGlyGluLysValValIleGlyArgMetSerAlaGlnVal 100
 DB 2365 GCTTCCATCCGAGACAGAGGAGAGGTGTCTGTATGGCCGATGACCCAGGTG 2306
 QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 DB 2305 CCCATGAACATGACCATCATCTGCTGCATGCTGCATCTTACAGGAAGACCCCAACCGTG 2246
 QY 121 ValPheTrpGlnThrValAlaAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 DB 2245 GTTCTGTGGCAGTGGGTGATCATGCTTCAATGCTTAACTACTCCAAACCGCACT 2186
 QY 141 GYASPTPrProIleThrValArgInLeuGlyThrAlaTyrValSerAlaThrThrGly 160
 DB 2185 GGTGACACTCCATACCTGTGAGCGACGCTGGGACAGCCTATGTGATGCCACCTGGA 2126
 QY 161 AlaValAlaThrAlaLeuGlyLeuIleuYsSerIleuThrLysHisLeuProProLeuValGly 180
 DB 2125 GCTGTGGCCAGCGCCCTGGGACTCAATCCCTCACCAAGACACTGGCCCCCTGGTGGCG 2066
 QY 181 ArgPheValProPheAlaAlaValAlaAlaAsnCysIleAsnIleProLeuMetArg 200
 DB 2065 AATTTTGTCTCCCTTTGACAGAGTGGGAGCTGCCAATGATCACTCCCTGTAGAGG 2006
 QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
 DB 2005 CAGAGAGAGCTCAGAGTGGGCAATCCCGGTGGCTGTAGAGCGAGCTGAGGCTTGCTAC 1946
 QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValAlaIleSerArgIleCysMetAla 240
 DB 1945 TCGGTGACTGCAGCAAGCAGGGAATCTCCAGTGTGATTTCAAGAAATCTGATGGCG 1886
 QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGluLysLysAspPheLeu 260
 DB 1885 ATTCTGCGATGGCCATCCACCACTGATCATGACACTGTGAGAAAGAAAGACTCTCG 1826
 QY 261 Lys 261
 DB 1825 AAG 1823
 RESULT 7
 ABL39805
 ID ABL39805 standard; cDNA; 984 BP.

XX ABL39805;
 AC 10-MAY-2002 (first entry)
 DE Human NS CDNA sequence SEQ ID NO:115.
 XX
 XX Human; cytosolic; osteopathic; gynaecological; neuroprotective;
 KM antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
 KM vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
 KM anorectic; muscular; antifertility; cardiovascular; anticoagulant;
 KM antifibrinolytic; hypotension; antiaesthetic; immunomodulator; cardiant;
 KM anticonvulsant; antidiabetic; tranquilizer; antidepressant; neuroleptic;
 KM gastrointestinal; virucide; antitumor; cerebroprotective; nootropic;
 KM contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
 KM endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 KM rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
 KM inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KM infertility; cardiovascular disease; coagulation disease; hypertension;
 KM ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
 KM diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KM gastric ulcer; Alzheimer's disease; gene; ss.
 OS Homo sapiens.
 XX
 XX WO200206315-A2.
 XX
 XX 24-JAN-2002.
 XX
 XX 17-JUL-2001; 2001WO-IL00653.
 XX
 XX 18-JUL-2000; 2000IL-0137345.
 XX
 XX 15-DEC-2000; 2000IL-0140354.
 XX
 XX (COMP-) COMPUGEN LTD.
 XX
 XX Mintz L, Freilich S, Bernstein J;
 XX
 XX WPI; 2002-155037/20.
 DR P-PSDB; ABB06151.
 XX
 XX One hundred and twenty eight novel nucleic acid sequences, useful for
 PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
 XX
 XX Claim 1: Page 141-142; 290pp; English.
 PS
 XX ABL39691 to ABL39818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
 CC (NS) can have cytosolic, osteopathic, gynaecological, neuroprotective,
 CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
 CC vasotropic, antiarteriosclerotic, antitumor, antiinflammatory, dermatological,
 CC anorectic, muscular, anti-HIV, antifertility, cardiovascular,
 CC anticoagulant, antifibrinolytic, hypotension, antiaesthetic, cardiant,
 CC immunomodulator, anticonvulsant, antidiabetic, tranquilizer, antitumor,
 CC antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
 CC nootropic and contraceptive activities. The NS can be used in vaccines,
 CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contraceptive.
 XX
 XX Sequence 984 BP; 224 A; 278 C; 285 G; 196 T; 1 other;
 SO
 Alignment Scores:
 Pred. No.: 6-77e-140 Length: 984
 Score: 1318.00 Matches: 260
 Percent Similarity: 99.62% Conservative: 0

Best Local Similarity: 99.62% Mismatches: 1
 Query Match: 97.99% Indels: 0
 DB: 24 Gaps: 0
 US-09-990-415a-2 (1-261) x ABL39805 (1-984)
 QY 1 MetGUseRlySMetGlyGluProLeuAAspIleAsnIleGlnGluProArgTrpAsp 20
 DB 49 ATGGAACCAAAATGGGGAATGGCTTTAGACATCAACATCCAGAACCTCGCGGAC 108
 QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
 DB 109 CAAGTACTTTCCTCGGAGAGC-CGGCACTTTTCACGTGTACGTATCGAATCTG 167
 QY 41 LeuLeuSerGlyAlaGlnLeuGlnLalaSerArgAsnIleValGlnAsnTyraArgAlaGly 60
 DB 168 CTGCTGTCGGGGGACACAGCTGGAACTTCTCGGAACATCGTCAGAACTACAGGGCCGC 227
 QY 61 ValValThrProGlyIleThrGluAspGlnLeuTrpArgAlaLysTyraValTyraPse 80
 DB 228 GTGGTGACCCCGAGGATCACCGAGACAGCTGTGAGGGCCCAAGTATGTATGACTCC 287
 QY 81 AlaPheHisProAspThrGlyGluLysValValIleuIleGlyArgMetSerAlaGlnVal 100
 DB 288 GCTTCCATCCGGACACAGGGAGAGGGTGTCTGATTTGGCCGATGTCACCCAGGTG 347
 QY 101 PrometAanMetThrIleThrGlyCysMetLeuThrPheTyraArgLysThrProThrVal 120
 DB 348 CCCATGAACATGACATCATCTGCTGCATGCTGCATCTTACAGGAAGCCCAACCGTG 407
 QY 121 ValPheTrpGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTySerAsnArgSer 140
 DB 408 GTGTTCTGGACGGTGGTGAATCATGCTTCAATGCATGTGTAACCTACCCAAACGCAAT 467
 QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyraValSerAlaThrThrGly 160
 DB 468 GTGACACTCCCATCTGTGAGGACGCGGGGACAGCTTATGTAGTGCACCACTGGA 527
 QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
 DB 528 GCTGTGGCCACGCGCCCTGGAGCTCAATCCCTCACCAAGACCTGCCCTTGGTCCGG 587
 QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
 DB 588 AGATTGTGCCCTTGGACAGAGTGGCAGCTGCCAATGCATCAATCAATCCCTGATGAGG 647
 QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTy 220
 DB 648 CAGAGAGAGCTGACAGTGGGCAATCCCGGTGTGTAGAGCGATCAGAGGCTTGACTTC 707
 QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValIleIleSerArgIleCysMetAla 240
 DB 708 TCGGTGACTGCAAGCAACAGAGGATCTTCCAGGTGTGTATTCAAAGATCGCATGGGG 767
 QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeu 260
 DB 768 ATTCCTGCATGAGCCATCCACACTGATCATGAGCACTGTGAGAGAAGAACTTCTCTG 827
 QY 261 Lys 261
 DB 828 AAG 830
 RESULT 8
 AAC75798 standard: cDNA; 843 BP.
 AAC75798:
 AC AAC75798;
 XX
 XX 08-FEB-2001 (first entry)
 DE Human ORF1353 polynucleotide sequence SEQ ID NO:2705.
 XX Human; open reading frame; ORF; detection; cytosolic; hepatotropic;
 KM

KW vulnary; antipariatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antirheumatic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antineuritic;
 KW antiviral; antibacterial; antifungal; antirheumatic; antihypertensive;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX
 XX Homo sapiens.
 OS
 PN MO200058473-A2.
 PD
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000MO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 PI Shinkets RA, Leach M;
 XX WPI: 2000-602362/57.
 DR P-PSDB; AAB41589.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 5; Page 1942; 5507pp; English.
 PS
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antipariatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antirheumatic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antineuritic; antineuritic; antiviral; antineuritic; antineuritic;
 CC antihypertensive; antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX
 SQ Sequence 843 BP; 183 A; 241 C; 252 G; 165 T; 2 other;
 Alignment Scores:
 Pred. No.: 2, 01e-134 Length: 843
 Score: 1269.00 Matches: 246
 Percent Similarity: 99.20% Conservative: 3
 Best Local Similarity: 98.01% Mismatches: 2
 Query Match: 94.35% Indels: 0
 DB: 21 Gaps: 0
 US-09-990-415a-2 (1-261) x AAC75798 (1-843)

QY 1 MetGluSerLysMetGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
 DB 91 ATGGAAGCAAAATGGGTGATATGCTTTAGACATCAACATCCAGAAACCTCGCTGGAGC 150
 QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
 DB 151 CAAAGTACTTTCTGGGCAAGAGCCCGCACTTTTCTACTGTTCATCTCGAATCTG 210
 QY 41 LeuLeuSerGlyAlaGlnLeuGlnAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 DB 211 CTGCTCTCCGGGGCACAGCTGGAAGCTTCTCGAATCATGTGAGAACTACAGAGCGCGG 270
 QY 61 ValValThrProGlyIleThrGluAspGlnLeuThrArgAlaLysTyrValTyrAspSer 80
 DB 271 GTGGTGACCCCAAGGATCAAGGAGGACGAGCTGTGAGGCGCAAGATGTGTGACTCC 330
 QY 81 AlaPheHisProAspThrGlyGluLysValIleLeuIleGlyArgMetSerAlaGlnVal 100
 DB 331 GCCTTCATCCGACACACAGGAGGAGGAGTGTCTGATTTGCCGCAATGACCCAGGTG 390
 QY 101 ProMetLysMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 DB 391 CCCATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATGAC 450
 QY 121 ValPheThrGlnThrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 DB 451 GTGTTCTGCAATGGGTGATCAAGCTTCAATGCAATGCAATGCAATGCAATGCAATGCA 510
 QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
 DB 511 GGTGACACATCCATGACATGACATGACATGACATGACATGACATGACATGACATGAC 570
 QY 161 AlaValAlaThrAlaLeuGlnLysSerLeuThrLysHisLeuProLeuValGly 180
 DB 571 GGTGACACATCCATGACATGACATGACATGACATGACATGACATGACATGACATGAC 630
 QY 181 ArgPheValProPheAlaIleValAlaIleAsnGlyIleAsnIleProLeuMetArg 200
 DB 631 AATTCGATACCTTTGACAGAGTGGAGCTGCACTGCACTGCACTGCACTGCACTGCACTG 690
 QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
 DB 691 CAGAGGAGCTGACAGTGGGATCCAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGA 750
 QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
 DB 751 TCGGTGACTGCTGCCAAACAGGCAATCTTCCAGTGTGTGTGATCGAATCGGATGCGCG 810
 QY 241 IleProAlaMetAlaIleProProLeuIleMet 251
 DB 811 ATCCCGCCATGCGCATTTCCCGGTGATCATG 843
 RESULT 9
 AAS84362
 ID AAS84362 standard; cDNA; 2215 BP.
 XX
 AC AAS84362;
 XX
 DT 13-FEB-2002 (first entry)
 DE
 DE DNA encoding novel human diagnostic protein #20166.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 PN MO200175067-A2.
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEO INC.
 PA Drmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 DR P-PSDB; ABG20175.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 1; SEQ ID NO 20166; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS6197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 2215 BP; 491 A; 649 C; 582 G; 492 T; 1 other:
 XX
 Alignment Scores:
 Pred. No.: 1,21e-128 Length: 2215
 Score: 1223.50 Matches: 256
 Percent Similarity: 89.82% Conservative: 0
 Best Local Similarity: 89.82% Mismatches: 2
 Query Match: 90.97% Indels: 27
 DB: 23 Gaps: 1
 US-09-990-415a-2 (1-261) x AAS84362 (1-2215)
 QY 4 LysMetGIyGIuLeuProLeuAspIleAsnIleGlnIuProArgTrpAspGlnSerThr 23
 DB 208 AAATGGGTGAATGGCTTGTAGACATCAACATCCAGGAACCTGGTGACCAAGTACT 267
 QY 24 PheLeuGIyARGAla-ArgHisPhePheThrValThrAspProArgAsnLeuLeuSe 43
 DB 268 TTCCTGGGCAGAGCGCGGCACTTTTACAGTACTATCTCTCAATTCGTCGCTC 327
 QY 43 rGIyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTrpArgAlaGlyValAlaThr 63
 DB 328 CGGGGCACAGCTGGAGAGCTTCTCGAATCATGTGAGAACTACAGGCGCGGTGTAC 387
 QY 63 rProGlyIleThrGlnAspGlnLeuTrpArgAlaLysTrpValTrpAspSerAlaPheH1 83
 DB 388 CCCAGGATACACCGAGGACAGCTGTGGAGGCGCCAGATGATGTGACTCCGCTTCCA 447
 QY 83 sProAspThrGIyGIuysValIleuIleGIyArgMetSerAlaGlnValPrometas 103
 DB 448 TCCGACACAGAGGGAGAAAGTGTGCTCATTTGGCGCATGTCAGCGCCAGGTGCCATGAA 507
 QY 103 mMetThrIleThrGIyCysMetLeuThrPheTr----- 114

DB 508 CATGACATCACTGCTGCATGCTCATCTTCTACAGGAGGGGTCTAAAGATAGAGGCCA 567
 QY 115 -----ArgLysThrProThrValValPh 122
 DB 568 CTGTAGACGGGGCAGACAGTACTGCTTGTTCCTCCCTCCAGAGACCCCAACCGGTGTT 627
 QY 122 eTrpGlnTrpValaGlnSerPheAsnAlaIleValaAsnTrpSerAsnArg-SerGIyA 142
 DB 628 CTGGCAGTGGGGATGATCAGTCTCTCAATGGCCATTTTAACTCACTCCACCCAGGTGGTGG 687
 QY 142 sPThr-ProIleThr-ValArgGlnLeuGIyThr-AlaTrpValSerAla-ThrThrGIy 160
 DB 688 ACACCTCCCATCACTGTGTAGAGCAGCTGGGAGCAGCCATATGTAGTGCANCCACTGGA 747
 QY 161 AlaValAlaThrAlaLeuGIyLeuysSerLeuThrLysIleuProLeuValGIy 180
 DB 748 GCTGTGGCCAGCGCCCTGGAGCTCAAAATCCCTCCACCAAGCAGCTCCCTTGTGTGGC 807
 QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnAsnIleAsnIleProLeuMetArg 200
 DB 868 AGATTGTGGCCCTTGTGACAGCAGTGGCAGCTGCCAATGCAATCATCCCTGATGAGG 867
 QY 201 GlnArgGluLeuGlnValGIyIleProValAlaAspGluAlaGIyGlnArgLeuGIyTr 220
 DB 868 CAGAGAGAGCTGCAGAGTGGGCAATCCCGGTGCTGTAGAGGAGCTCAAGGCTTGCTAC 927
 QY 221 SerValThrAlaAlaLysGlnGIyIlePheGlnValIleSerArgIleCysMetAla 240
 DB 928 TCGGTGACTGCAGCCACAGGAGAAATCTTCCAGGTGTGATTTCTAGAAATCTGCATGGCG 987
 QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysAspPheLeu 260
 DB 988 ATTCTGTGCATGGCCATCTCCACACTGATCATGAGACACTCTGTGAGAAAGAACTTCCG 1047
 QY 261 Lys 261
 DB 1048 AAG 1050
 RESULT 10
 ID AAI59569 standard; cDNA; 1076 BP.
 XX AAI59569;
 AC AAI59569;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 1772.
 XX
 KW Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW Leukaemia; ss.
 XX
 OS Homo sapiens.
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0596042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX

(HXSE-) HXSEQ INC.

PA Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR P-PSDB: AAM40413.

PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 XX
 PS Claim 1, SEQ ID NO 1772; 10078bp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 CC
 SQ Sequence 1076 BP; 285 A; 258 C; 269 G; 264 T; 0 other.

Alignment Scores:
 Pred. No.: 8.2e-114 Length: 1076
 Score: 1089.00 Matches: 201
 Percent Similarity: 89.45% Conservative: 28
 Best Local Similarity: 78.52% Mismatches: 27
 Query Match: 80.97% Indels: 0
 DB: 22 Gaps: 0

US-09-990-415a-2 (1-261) x AA159569 (1-1076)

QY 6 GlycylleuProleuAspIleAsnIleGlnIupProargTrpAspGlnSerThrPheIeu 25
 DB 67 GGAGAACTACCAACCAATTAACATCAAGAACTCGATCGATCAAGCACTTCAAT 126
 QY 26 GlyArgAlaArgHisPhePheThrValThrAspProArgAsnIleuLeuSerGlyAla 45
 DB 127 GGACGAGCCATCATTTCTTCACTGTAACCTGACCCGAGCAATTCCTGTAACCAACAA 186
 QY 46 GlnIeuGlnAlaSerArgAsnIleValGlnAsnTrpArgAlaGlyValThrProGly 65
 DB 187 CAACTCGAGAGTGGAGAAATAAGTACATGATTAACAGGCAAGATTCCTCTCGT 246
 QY 66 IleThrGluAspGlnIeuTrpArgAlaIleValTrpAspSerAlaPheHisProasp 85
 DB 247 CTTCACGAATAATGATTTGTGAGAGCAACAGTACATCAATGATCACTTTCATCTGAC 306
 QY 86 ThrGlyGlnLysValIleValIleGlyArgMetSerAlaGlnValProMetAsnMetThr 105
 DB 307 ACTGGTGAGAGATGATTTGATAGAGAAATGTCAGCCAGTCCCATGACATGAC 366
 QY 106 IleThrGlyCysMetLeuThrPheThrArgGlyThrProThrValValPheTrpGlnTrp 125
 DB 367 ATCAAGAGTTTATGATACGCTTTACAGACTACGCGCTGCTCTGTCGAGTGG 426
 QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTrpSerAsnArgSerGlyAspThrProIle 145
 DB 427 ATTAACAGCTCTCAATGCGCTGCTCAATTAACACCAAGAGTGGAGCGACCCCTC 486
 QY 146 ThrValArgGlnIeuGlyThrAlaIleValSerAlaThrThrGlyAlaValAlaThrAla 165

DB 487 ACTGCAATGAGTTGGGAGACAGCTTACCTTTCTGCACAACTGTCCTGACCAAGACT 546
 QY 166 LeuGlyLeuLysSerIeuThrLysHisIleuProleuValGlyArgPheValProPhe 185
 DB 547 CTAGAGACTCAATGATTCATCCAGAGATCTTCACACAGATGAGAGCGTTTGTCCCTT 606
 QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIlePheLeuMetArgGlnArgGln 205
 DB 607 GCTGCCGTAGCTGCTGCTAATTCATTAATTCATTAATGAGCAAGGCAAGCTCAA 666
 QY 206 ValGlyIleProValAlaAspGlnValIleGlnArgGlyIleValSerValThrAlaIle 225
 DB 667 GTTGCAATTCCTCCGTCACGATGAGAAATGGAAACCGCTTGGGGAAGTCCGCAAGCTTCG 726
 QY 226 LysGlnGlyIlePheGlnValIleSerArgIleCysMetAlaIleProAlaMetAla 245
 DB 727 AAACAGGACATCAGCAGAGTGTGCTGTCAGAGATTCATGATGAGCCCTGGCATGGCC 786
 QY 246 IleProProLeuIleMetAspThrIeuGlnLysLysAspPheLeuLys 261
 DB 787 ATCCCTCATTCATTAATGAAACACTTGGAAAGAAAGCCCTTTTGAAG 834

RESULT 11
 ABN84365
 ID ABN84365 standard; cDNA: 1134 BP.
 AC ABN84365;
 XX
 XX 01-OCT-2002 (first entry)
 DE
 DE Alzheimer's disease associated protein ADPI-41 coding sequence.
 XX
 KW ADPI-41; human; neuropsychiatric disorder; neurological disorder;
 KW Alzheimer's disease; nootropic; neuroprotective; neuroleptic;
 KW antiparkinsonian; antidepressant; diagnosis; vaccine; human;
 KW sideroflexin 1; sfxn1; gene therapy; gene; ss.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 11..979
 FT /tag= a
 FT /product= "ADPI-41"

W0200246221-42.
 PD 13-JUN-2002.
 XX
 XX 10-DEC-2001; 2001WO-GB05459.
 XX
 XX 08-DEC-2000; 2000US-254431P.
 PR (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
 PA
 PI Herath HMC, Parekh RB, Rohlf C, Terrett JA, Tyson KL;
 DR WPI: 2002-557575/59.
 DR P-PSDB: ABB79572.
 XX
 XX Isolated or recombinant polypeptide useful in the treatment of
 PT neuropsychiatric or neurological disorder comprises a specific amino
 PT acid sequence -
 XX
 XX Claim 6; Fig 2a; 51pp; English.

CC The present sequence is the coding sequence of a human protein,
 CC named ADPI-41, which is differentially expressed (present in
 CC reduced amounts) in all areas of the brain tested except the
 CC amygdala in Alzheimer's disease (AD). The full-length open
 CC reading frame was amplified by PCR from human brain and liver
 CC cDNA libraries. ADPI-41 is useful as a marker and/or therapeutic
 CC target in AD. Its amino acid sequence is 95% identical to that of
 CC murine sideroflexin 1 (Sxfln1), recently identified in the mouse

CC model of sideroblastic anemia. This is the first demonstration of
 CC the human Sxtn protein and the first indication that it may be
 CC associated with the occurrence of AD. ADPI-41 undergoes
 CC alternative splicing to result in an isoform of 261 amino acids
 CC (see AB079573). The invention provides ADPI-41 polypeptides and
 CC polynucleotides, vectors, host cells and antibodies. A claimed
 CC method of screening for and/or diagnosis of a neuropsychiatric or
 CC neurological disorder in a subject and/or monitoring the
 CC effectiveness of therapy comprises detecting and/or quantifying
 CC the amount of ADPI-41 polypeptide and/or nucleic acid molecule in a
 CC sample from the subject. A claimed method for the prophylaxis
 CC and/or treatment of a neuropsychiatric disorder or neurological
 CC disorder comprises administering an ADPI-41 polypeptide or nucleic
 CC acid. A method of screening for agents that modulate ADPI-41
 CC activity or expression is also claimed. Disorders that may be
 CC treated or diagnosed include AD, vascular dementia, Lewy body
 CC dementia, schizophrenia, Parkinson's disease, multiple sclerosis,
 CC and depression.

XX Sequence 1134 BP; 303 A; 268 C; 273 G; 287 T; 3 other;

Alignment Scores:

Pred. No.: 8 85e-114 Length: 1134
 Score: 1089.00 Matches: 201
 Percent Similarity: 89.45% Conservative: 28
 Best Local Similarity: 78.52% Mismatches: 27
 Query Match: 80.97% Indels: 0
 DB: 24 Gaps: 0

US-09-990-415a-2 (1-261) x ABN04365 (1-1134)

QY 6 GlycyluLeuProLeuAspIleAsnIleGlnGluProArgTrpPaSpGlnSerThrPheLeu 25
 DB 17 GGAAGAACTACCAACCAATTAACTACAGAGCACTCGATGGATCAAGACCTTCATT 76
 QY 26 GlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
 DB 77 GGAGAGCCCAATCTTCTCTACGTACCTGACCCCGAGCAATCTGTTAAACCAAGAA 136
 QY 46 GlnLeuGlnAspSerArgAsnIleValGlnAsnTrpArgAlaValValThrProGly 65
 DB 137 CAACCTGAGAGTGGGAGAAAATAGTACATGATTCACAGCAGCAATTTGTCCTGGT 196
 QY 66 IleThrGlnAspGlnLeuTrpArgAlaLysTrpValTrpAspSerAlaPheHisProAsp 85
 DB 197 CTTCACGAAATATGATTTGTGAGAGCAAGATCATCTATGATTCAGCTTTTCATCTGAC 256
 QY 86 ThrGlyGlnLysValValIleGlyArgMetSerAlaGlnValPrometAsnMetThr 105
 DB 257 ACTGTGAGAAAGATGATTTGATNGAGAAATGTCACGCCAGGTTCATGAAATGAC 316
 QY 106 IleThrGlyCysMetLeuThrPheTrpArgLysThrProThrValValPheTrpGlnTyr 125
 DB 317 ATCAGAGCTGTATGATGAGAGCTTTACAGACATCGCGCTGCTGCTCTGCACTGG 376
 QY 126 ValAsnGlnSerPheAsnAlaIleValAsnTrpSerAsnArgSerGlyAspThrProIle 145
 DB 377 ATTAAACAGTCTTCATGCGCTGCTGATTCACCAAGAAATGAGAGCAGCACCCCTC 436
 QY 146 ThrValArgGlnLeuGlyThrAlaTrpValSerAlaThrArgValAlaValAlaThrAla 165
 DB 437 ACTGTCAATGAGTTGGAGACAGCTTACGTTTCTGCAACCAATGCTGCGTAGCAACAGCT 496
 QY 166 LeuGlyLeuSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
 DB 497 CTAGAGCTCAATGATGACCAAGCATGCTCACACCACTGATAGAGCTTTGTTCCCTTT 556
 QY 186 AlaAlaValAlaAlaAsnGlyIleAsnIleProLeuMetArgGlnArgGlnLeuGln 205
 DB 557 GCTGCGCTAGCTGCTGATTAATGATTAATTCATTAATGAGCAAGGAAAGCACTGAAA 616
 QY 206 ValGlyIleProValAlaAspGlnValGlnArgLeuGlyTrpSerValThrAlaAla 225

DB 617 GTTGCCATTCCTCCGTCACGATGAGTAAGGAAACCGCTTTGGGAGCTGGCAACGCTGCG 676
 QY 226 LysGlnGlyIlePheGlnValValIleSerArgIleCysMetAlaIleProAlaMetAla 245
 DB 677 AAACAAGCAATCAGCAAGTGTCTGTCAAGATTCATGCGACGCCCTGGCAATGGCC 736
 QY 246 IleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeuLys 261
 DB 737 ATCCCTCATTCATTATTAACACACTTGGAAAAAGAACGCTTTTGAAG 784

RESULT 12

AL26555 standard; cDNA: 1485 BP.

AC AL26555;

DI 07-DEC-2001 (first entry)

DE Human breast cancer expressed polynucleotide 19012.

KW Human; breast cancer; cell marker; cytostatic; ss.

OS Homo sapiens.

PN WO200151628-A2.

PD 19-JUL-2001.

PF 10-JAN-2001; 2001WO-US00798.

PR 14-JAN-2000; 2000US-0176077.

PR 14-MAR-2000; 2000US-0189167.

PR 24-MAR-2000; 2000US-0192099.

PR 29-MAR-2000; 2000US-0193480.

PR 15-MAY-2000; 2000US-0205230.

PR 09-JUN-2000; 2000US-0211315.

PR 25-JUL-2000; 2000US-0220534.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Lillie J, Xu Y, Wang Y, Steinmann K;

DR WPI: 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer

PS Claim 1; Page 3556; 3655pp; English.

XX The invention relates to human breast cancer expressed polynucleotides

CC (AL07544-AL26789) and methods of assessing whether a patient is

CC afflicted with breast cancer by examining the correlation between the

CC expression of certain markers and the cancerous state of breast cells.

CC The polynucleotides and encoded polypeptides are potential markers for

CC detecting, diagnosing, monitoring, characterizing, treating and

CC potentially preventing breast cancer. The polynucleotides and encoded

CC polypeptides are also useful for isolating compounds with cytostatic

CC activity.

SO Sequence 1485 BP; 375 A; 378 C; 365 G; 359 T; 8 other;

Alignment Scores:

Pred. No.: 1.31e-113 Length: 1485
 Score: 1089.00 Matches: 201
 Percent Similarity: 89.45% Conservative: 28
 Best Local Similarity: 78.52% Mismatches: 27
 Query Match: 80.97% Indels: 0
 DB: 22 Gaps: 0

US-09-990-415a-2 (1-261) x AL26555 (1-1485)

QY 6 GlycyluLeuProLeuAspIleAsnIleGlnGluProArgTrpPaSpGlnSerThrPheLeu 25
 DB 112 GGAAGAACTACCAACCAATTAACTACAGAGCACTCGATGGATCAAGACCTTTCATT 171

```

OY 26 GYARGAIAARGHISPhetherValThrasProARgSnleuLeuSerGlyala 45
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172 GACACAGCCAAATCATTTCTTCACTGTAAGTACCCAGAGAAATTTGTAAACCAAGAA 231
OY 46 GlnleuGluAlaSerARgAsnIleValGlnAsnTYrARgAlaGlyValValThrProGly 65
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 222 CAACTCGAGAGTGCAGAAAATAATAGTACATGATTACAGCAAGAAATTTCTCTCTGT 291
OY 66 IlerhrgluAspGlnleuThrPARgAlaLysTYrValTYrAspSerAlaPheHisProasp 85
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 292 CTTACAGAAAATGAATGTGTGAGAGCAAAAGTACACTATGATTGACTTTTATCTCTGAC 351
OY 86 ThrGlyGluLysValValleuIleGlyARgMetSerAlaGlnValProMetAsnMetThr 105
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 352 ACTGCTGAGAAATGATTTTGTATGAGAAATGTCAACCCAGAGTCCCATGAAACATGACC 411
OY 106 IlerhrglyCyMetleuThrPheTYrARgLYrThrProThrValValPheThrPglInTP 125
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 412 ATCAGAGGTTGTATGATGACGTTTACAGACTACAGCCGCGTGTGTTCTGGCAGTGG 471
OY 126 ValaIngInSerPheAsnAlaIleValAsnTYrSerAsnARgSerGlyAspThrProIle 145
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 472 ATTTAACCAAGTCTTCAATGCCCTGCTCAATTACACCAAGAAAGTGAAGCCACCCCTTC 531
OY 146 ThrValARgGlnleuGlyThralATyValSerAlaThrThrGlyAlaValAlaThrAla 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 532 ACTGTCATGATGATTGGGAACACCTTACGTTTCTGCAACACAGTGAGCCCTAGCAACAGCT 591
OY 166 IeuglyleuLysSerleuThrLysHisleuProPoleuValGlyARgPheValProPhe 185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 592 CTAGAGACTCAATGCACTTGACCAAGCATGCTCACCACATGATGAGAGCTTTGTCCTCTT 651
OY 186 AlaAlaValAlaAlaAlaAsnCySileAsnIleProleuMetARgGlnleuGln 205
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 652 GCTGCCGATGCTGCTGCTAATGCAATTAATTCATTATATAGCCAAAGGAACTCAAA 711
OY 206 ValGlyIleProValAlaAlaAspGluAlaGlyGlnARgLeuGlyTYrSerValThrAla 225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 712 GTTGGCATTTCCCGTCAAGCATGATGAAATGGAAACCGTTGGGGAGTGGCCGAACGCTGCG 771
OY 226 LysGlnGlyIlePheGlnValValIleSerARgIleCyMetAlaIleProAlaMetAla 245
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 772 AAACAGCCATCAGCAAGATTGTCGTGTCAGAGATTCTCATGCGAGCCCTGCGCATGGCC 831
OY 246 IleProPoleuIleMetAspThrleuGlnLysLysAspPheLeuLys 261
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 832 ATCCCTCATTCATTATGAACACTTGGAAAAAGAAAGCCCTTTTGAAG 879

```

RESULT 13

AA161354
ID AA161354 standard; cDNA; 2053 BP.

AC AA161354;

DT 22-OCT-2001 (first entry)

DE Human polynucleotide SEQ ID NO 5343.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokine; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.

OS Homo sapiens.

PN WO200153312-A1.

XX 26-JUL-2001.

PF 26-DEC-2000; 2000MO-US34263.

```

XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0596042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao Q, Zhou P, Goodrich R, Dymnac RT;
XX
XX WPI: 2001-442253/47.
XX P-PSDB; AAM42198.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 5343; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cyostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokine activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening
XX assays for receptor activity, arthritis and inflammation, leukaemia and
XX C.N.S disorders.
XX
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
XX SQ Sequence 2053 BP; 582 A; 436 C; 451 G; 584 T; 0 other;
XX
XX Alignment scores:
XX Pred. No.: 2,1e-113 Length: 2053
XX Score: 1089.00 Matches: 201
XX Percent Similarity: 89.45% Conservative: 28
XX Best Local Similarity: 78.52% Mismatches: 27
XX Query Match: 80.97% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-990-415a-2 (1-261) x AA161354 (1-2053)
OY 6 GlyGluLeuProLeuAspIleAsnIleGlnGluProARgTrPARspGlnSerThrPheLeu 25
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 94 GGAGAACTACACCAACAACATTAACAAGAACTCGATGGAGATCAAAAGCACTTCAT 153
OY 26 GYARGAIAARGHISPhetherValThrasProARgAsnleuLeuSerGlyala 45
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 154 GCACGAGCCAAATCATTTCTTCACTGTAAGTACCCAGAGAAATTTCTGTAACCAAGAA 213
OY 46 GlnleuGluAlaSerARgAsnIleValGlnAsnTYrARgAlaGlyValValThrProGly 65
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 214 CAACTCGAGAGTGCAGAAAATAATAGTACATGATTACAGCAAGAAATTTCTCTCTGT 273
OY 66 IlerhrgluAspGlnleuThrPARgAlaLysTYrValTYrAspSerAlaPheHisProasp 85
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 274 CTTACAGAAAATGAATGTGTGAGAGCAAAAGTACATGATTGATTGACTTTTATCTCTGAC 333
OY 86 ThrGlyGluLysValValleuIleGlyARgMetSerAlaGlnValProMetAsnMetThr 105
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 334 ACTGCTGAGAAATGATTTTGTATGAGAAATGTCAAGCCAGGTTCCCATGAAACATGACC 393

```

QY 106 IlethrglyCysMetLeuThrPheTyArgIysThrProThrValValPheTrpGlnTrp 125
 DB 394 ATCACAGGTGTATGATGACGTTTACAGACTACGCGCGGTGCTGTCTGGCAGTGG 453
 QY 126 ValangInserPheAsnAlaIleValAsnTySerAsnArgSerGlyAspThrProIle 145
 DB 454 ATTAACGAGTCCTTAATGCGGTCTCATTAACCAACAAGAGGAGACGACCCCTGC 513
 QY 146 ThrValArgInLeuGlyThrAlaTyValSerAlaThrThrGlyAlaValAlaThrAla 165
 DB 514 ACTGTCATGAGTGGGAAACAGCTTACCTTCTGCAACAACGTGGCGGTAGCAACACT 573
 QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
 DB 574 CTAGACACTCAATGCAATGACCAAGCATCTCCACCAAGATGAGGAGCGTTTGTCCCTTT 633
 QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlnLeuGln 205
 DB 634 GCTGCGTAGCTGCTGCTAATGCTAATTAATTCATTAAATGAGCAAGGAACTCAAA 693
 QY 206 ValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTySerValThrAlaAla 225
 DB 694 GTTGCCATTCCTCCGTCACGATGAGAAATGGAACCGCTTGGGGAGTGGCGAAGCTCG 753
 QY 226 LysGlnGlyIlePheGlnValAlaIleSerArgIleCysMetAlaIleProAlaMetAla 245
 DB 754 AAACAGGCAATCAGCAAGTGTGTGTCAGGATTCATCAATGAGCAAGCGCCCTGGCATG 813
 QY 246 IleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeuLys 261
 DB 814 ATCCCTCATTCATTATGACACCTTTGGAAAAAGAAAGCCCTTTTGAAG 861

RESULT 14

ABN83757 standard; cDNA: 2269 BP.

ABN83757;

02-SEP-2002 (first entry)

Human protein cluster I gene.

Protein cluster I: human; metabolic disorder; obesity; diabetes;

antidiabetic; diagnosis; therapy; gene; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 125..1093

FT /tag= a /product= "Protein cluster I"

MO200242324-A1.

30-MAY-2002.

22-NOV-2001; 2001WO-SE02581.

24-NOV-2000; 2000SE-0004325.

(PHAA) PHARMACIA AB.

Attersand A;

WPI; 2002-500277/53.

P-PSDB; ABB76448.

Novel nucleic acid molecule encoding protein cluster I, useful in the
 diagnosis of metabolic diseases, such as obesity and diabetes, and in
 the identification of agents useful in the treatment of the diseases
 Claim 1; Page 30-32; 34pp; English.

CC The present invention relates to the identification of a human gene
 CC family (see ABN83754-57) encoding a group of polypeptides (see
 CC ABN76445-48) referred to as Protein Cluster I. This family of
 CC homologous proteins was identified by an "all-versus-all" BLAST
 CC procedure using all Caenorhabditis elegans proteins in the
 CC Wormpep20 database release. Proteins of unknown function were
 CC compared to the Drosophila melanogaster Flybase database, and
 CC non-annotated protein clusters, conserved in both C. elegans and D.
 CC melanogaster, were used in a BLAST procedure against the Celera
 CC Human Genome Database, and Protein Cluster I proteins of unknown
 CC function were selected for study. The human part of Protein
 CC cluster I comprises polypeptides encoded by 3 genes. The present
 CC gene is expressed primarily in the immune system. The claimed
 CC protein cluster I nucleic acid molecules and proteins are proposed
 CC to be useful for differential identification of the tissue(s) or
 CC cell type(s) present in a biological sample, for diagnosis of
 CC diseases and disorders, including metabolic disorders and immune
 CC disorders, especially obesity and diabetes, and for identifying
 CC agents useful in the treatment of such diseases. The nucleic acid
 CC molecules are also useful as hybridisation probes, for chromosome
 CC and gene mapping, in PCR technologies, and in the production of
 CC sense or antisense nucleic acids.

SQ Sequence 2269 BP, 631 A; 479 C; 506 G; 652 T; 1 other:

Alignment Scores:

Pred. No.:	2,42e-113	Length:	2269
Score:	1089.00	Matches:	201
Percent Similarity:	89.45%	Conservative:	28
Best Local Similarity:	78.52%	Mismatches:	27
Query Match:	80.97%	Indels:	0
DB:	24	Gaps:	0

US-09-990-415A-2 (1-261) x ABN83757 (1-2269)

QY 6 GlyGluLeuProLeuAspIleAsnIleGlnGluProArgTrpAsnGlnSerThrPheLeu 25
 DB 131 GGAGAACTTACCAACCAATTAACATCAAGGAACCTGATGATCAAGCAAGCACTTCACT 190
 QY 26 GlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeuLeuSerGlyAla 45
 DB 191 GGACAGGCAATCATTTCTTCACTTAAGTACAGCCAGCAAGCAATCTTTAAACCAAGAA 250
 QY 46 GlnLeuGluAlaSerArgAsnIleValGlnAsnTyArgAlaGlyAlaValThrProGly 65
 DB 251 CAACTCGAGATGCGAGAAATAAGTACATGATTAAGGAGGAGGAATGTCCTCTGCT 310
 QY 66 IlethrgIuAspGlnLeuTrpArgAlaLysTyValTyArgAspSerAlaPheHisProasp 85
 DB 311 CTACAGAAATGAATTTGTGGAGCAAGTACATCTATGATTCAGCTTTTCATCTGAC 370
 QY 86 ThrGlyGluLysValAlaLeuIleGlyArgMetSerAlaGlnValProMetAsnMetThr 105
 DB 371 ACTGCTGAAGATGATTTTGATAGGAAGATGTAGCCAGCTCCCATGAACATGAC 430
 QY 106 IlethrgIyCysMetLeuThrPheTyArgIysThrProThrValValPheTrpGlnTrp 125
 DB 431 ATCACAGGTGTATGATGACGTTTACAGACTACGCGCGGTGCTGTCTGGCAGTGG 490
 QY 126 ValangInserPheAsnAlaIleValAsnTySerAsnArgSerGlyAspThrProIle 145
 DB 491 ATTAACGAGTCCTTAATGCGGTCTCATTAACCAACAAGAGGAGACGACCCCTGC 550
 QY 146 ThrValArgInLeuGlyThrAlaTyValSerAlaThrThrGlyAlaValAlaThrAla 165
 DB 551 ACTGTCATGAGTGGGAAACAGCTTACCTTCTGCAACAACGTGGCGGTAGCAACACT 610
 QY 166 LeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGlyArgPheValProPhe 185
 DB 611 CTAGACACTCAATGCAATGACCAAGCATCTCCACCAAGATGAGGAGCGTTTGTCCCTTT 670
 QY 186 AlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArgGlnArgGlnLeuGln 205

QY 226 lysGlnGlyIlePheGlnValIleSerArgIleCysMetAlaIleProAlaMetAla 245
 ||||| ||| |||||:::||||| ||||| ||| |||||
 Db 791 AAACAAGCCATCAGCGAGTTGCGTCCAGCATTCATGCGCAGCCCCCTGCGCATGCC 850
 QY 246 IleProPheLeuIleMetAspThrLeuGluLysLysAspPheLeuLys 261
 ||||| |||||:::||||| ||||| ||| |||||
 Db 851 ATCCCTCCATTCATTATGAACACTTTGAAAGAAAGCCCTTTTGAAG 898

Search completed: October 16, 2003, 16:10:52
 Job time : 270 secs

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 US-09-990-415a-2 (1-261) x US-09-990-415a-1 (1-1232)

QY 1 MetGlnSerlysmetGlyGluLeuProLeuAspIleasnIleGlnGluProArgTTPasp 20
 DB 450 ATGGAAGCAAAATGGGTGAATTCCTTTAGACTCAATCCAGGAACCTCGCTGGAGAC 509

QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
 DB 510 CAAGAGTACTTCTCGGACAGACCCGGCACTTTTCTGCTGATGATCCCGAAATCTG 569

QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 DB 570 CTGCTGTCCGGGCGACAGTGGAGAGTCTTCGGAACTCGTGAGAACTAGAGCCGGCG 629

QY 61 ValValThrProGlyIleThrGlyAspGlnLeuThrArgAlaAlaGlyTyrValTyrAspSer 80
 DB 630 GTGGTACCCAGGAGATCCAGGAGACAGTGGAGGCGCAAGTGTGTATGACTCC 689

QY 81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
 DB 690 GCCTTCATCCGACACAGGAGGAGAGAGTGTCTGCTGATGCGCATGTACGCCAGGTG 749

QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 DB 750 CCCATGAACATGACATCACTGCTGCTCATCATTCATACAGGAAGACCCCAACCTG 809

QY 121 ValPheThrGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 DB 810 GTTCTTGGCAGTGGGTGAATTCCTTCAATGCCATTTGTAACTCACTCAACCCGAGT 869

QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
 DB 870 GGTGACACTCCATCACTGAGGAGAGTGGGACAGCCATGTGAGTGGCACCACTGGA 929

QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProLeuValGly 180
 DB 930 GGTGTGGCCAGGCGCTGGGACCTCAATCCCTCACCAAGACCTGCCCTTGTCGGC 989

QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
 DB 990 AGATTGTGCTTTCGACAGAGTGGCACTGCCAATGCAATCCCTGATGAGG 1049

QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
 DB 1050 CAGAGAGAGCTGCAGGTGGGCACTCCGGTGGTGTGAGAGAGAGTGCAGAGCTTGGCTAG 1109

QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
 DB 1110 TCGGTACTGCAACCAAGCAGGGAATCTTCAGGTGTGTCAAGAACTCCTCAGGCGC 1169

QY 241 IleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlnLysLysAspPheLeu 260
 DB 1170 ATTCCGCGCATGGCCATCCACCACTGATCATGAGACCTGTGAGGAAGAAAGACTTCCTG 1229

QY 261 Lys 261
 DB 1230 AAG 1232

RESULT 2
 US-09-867-550-1343
 Sequence 1343, Application US/09867550
 Patent No. US20020082206A1
 GENERAL INFORMATION:
 APPLICANT: Leach, Martin D.
 APPLICANT: Mehraban, Foad
 APPLICANT: Conley, Pamela
 APPLICANT: Law, Debbie
 APPLICANT: Topper, James
 TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and

TITLE OF INVENTION: Thereby
 FILE REFERENCE: 21402-013 (Cura-313)
 CURRENT APPLICATION NUMBER: US/09/867,550
 CURRENT FILING DATE: 2001-09-20
 PRIOR APPLICATION NUMBER: USSN 60/208,427
 PRIOR FILING DATE: 2000-05-30
 NUMBER OF SEQ ID NOS: 2125
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1343
 LENGTH: 1375
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: (1)..(2)
 OTHER INFORMATION: wherein any n is one of a or t or c or g
 US-09-867-550-1343

Alignment Scores:
 Pred. No.: 3,79e-171 Length: 1375
 Score: 1345.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-990-415a-2 (1-261) x US-09-867-550-1343 (1-1375)

QY 1 MetGlnSerlysmetGlyGluLeuProLeuAspIleasnIleGlnGluProArgTTPasp 20
 DB 91 ATGGAAGCAAAATGGGTGAATTCCTTTAGACTCAATCCAGGAACCTCGCTGGAGAC 150

QY 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
 DB 151 CAAGAGTACTTCTCGGACAGACCCGGCACTTTTCTGCTGATGATCCCGAAATCTG 210

QY 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
 DB 211 CTGCTGTCCGGGCGACAGTGGAGAGTCTTCGGAACTCGTGAGAACTACAGAGCCGGC 270

QY 61 ValValThrProGlyIleThrGlyAspGlnLeuThrArgAlaAlaGlyTyrValTyrAspSer 80
 DB 271 GTGGTACCCAGGAGATCCAGGAGACAGTGGGAGCCAAAGTGTGTATGACTCC 330

QY 81 AlaPheHisProAspThrGlyGluLysValValLeuIleGlyArgMetSerAlaGlnVal 100
 DB 331 GCCTTCATCCGACACAGGAGGAGAGGTGCTCCTGATTTGGCGCATGTCAGCCAGGTG 390

QY 101 PrometAsnMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
 DB 391 CCCATGAACATGACATCACTGCTGCTCATGCTCAATTCATACAGGAAGACCCCAACCTG 450

QY 121 ValPheThrGlnTrpValAsnGlnSerPheAsnAlaIleValAsnTyrSerAsnArgSer 140
 DB 451 GTTCTTGGCAGTGGGTGAATTCCTTCAATGCCATTTGTAACTCACTCAACCCAGT 510

QY 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
 DB 511 GGTGACACTCCATCACTGAGGAGAGTGGGAGAGCCATGATGTAGTGCACCACTGGA 570

QY 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
 DB 571 GGTGTGGCCAGGCGCTGGGACCAATCCCTCACCAAGCAGTGGCCCTTGTCGGC 630

QY 181 ArgPheValProPheAlaAlaValAlaAlaAlaAsnCysIleAsnIleProLeuMetArg 200
 DB 631 AGATTGTGCTTTCGACAGAGTGGCACTGCCAATGCAATCCCTGATGAGG 690

QY 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
 DB 691 CAGAGAGAGCTGCAGGTGGGCACTCCGGTGGTGTGAGAGGAGTGCAGAGGCTTGGCTAG 750

QY 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240

```
Db 751 TCCTGACCTGCACCAAGCAGGAGATCTTCAGGTGATTCACAAATCTGCATGGCG 810
Qy 241 ILleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlyLysAspPheLeu 260
Db 811 ATTCTGCTCCATGGCGCATCCACCACTGATCATGTGACACTCTGTGGAGAAAGAACTTCTG 870
Qy 261 Lys 261
Db 871 AAG 873

RESULT 3
US-10-037-270-1016
; Sequence 1016, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Duntui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Dermanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/10/037, 270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: PL_FL_genes Version 1.0
; SEQ ID NO 1016
; LENGTH: 1729
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (247)..(1371)
US-10-037-270-1016

Alignment Scores:
Pred. No.: 5,45e-171 Length: 1729
Score: 1345.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-990-415a-2 (1-261) x US-10-037-270-1016 (1-1729)

Qy 1 MetGluSerLysMetGlyGluLeuProLeuAspIleAsnIleGlnGluProArgTyrPasp 20
Db 247 ATGGAAGCAAAATGGGTGAATGCTTGTAGACATCAACATCCAGAACTCGCTGGGAC 306
Qy 21 GlnSerThrPheLeuGlyArgAlaArgHisPhePheThrValThrAspProArgAsnLeu 40
Db 307 CAAAGTACTTCTCTGGGCAAGCCGCGACATTTTTCATCTGTACTGATCTCCGAATCTG 366
Qy 41 LeuLeuSerGlyAlaGlnLeuGluAlaSerArgAsnIleValGlnAsnTyrArgAlaGly 60
```

```
Db 367 CTGCTGCCGGGGCACAGCTGGAAGCTTCTCGGAACATCGTCAGAACTACAGGCCGGC 426
Qy 61 ValValThrProGlyIleThrGluAspGlnLeuThrPargAlaLysTyrValTyrAspSer 80
Db 427 GTGGTACCCCAAGGATACCGAGACCACTGTGGAGGCCAAGTATGTGTATGATCC 486
Qy 81 AlaPheHisProAspThrGlyGlyLysValValLeuIleGlyArgMetSerAlaGlnVal 100
Db 487 GCCTTCATCCGGACACAGGGAGAAAGGTGGTCTGATTTGGCCGACATGACGCCAAGTG 546
Qy 101 PrometAspMetThrIleThrGlyCysMetLeuThrPheTyrArgLysThrProThrVal 120
Db 547 CCCATGAACATGACCAACACTGCTGCTGATCTCATATTCACAGGAAGACCCCAACCTG 606
Qy 121 ValPheThrGlnThrValAsnGlnSerPheAsnAlaIleValAsnTyrSerAspArgSer 140
Db 607 GTGTTCTGGCAGTGGGTGAATCACTCTTAATGCCATTTGTTAACTACTCCACCGCAGT 666
Qy 141 GlyAspThrProIleThrValArgGlnLeuGlyThrAlaTyrValSerAlaThrThrGly 160
Db 667 GGTGACACTCCCATCACTGTGAGGACACTGGGACAGCCATATGTGAGTGCACACATGGA 726
Qy 161 AlaValAlaThrAlaLeuGlyLeuLysSerLeuThrLysHisLeuProProLeuValGly 180
Db 727 GCTGTGGCACAGGCCCTGGGACCTCAATCCCTCACCAAGACACCTGCCCTGTGGTGGC 786
Qy 181 ArgPheValProPheAlaAlaValAlaAlaAlaAspCysIleAsnIleProLeuMetArg 200
Db 787 AGATTGTGCCCTTTGACAGAGTGGCAGCTGCCACTGCATCAACATCCCCCTATAGG 846
Qy 201 GlnArgGluLeuGlnValGlyIleProValAlaAspGluAlaGlyGlnArgLeuGlyTyr 220
Db 847 CAGAGACAGCTGCAGGTGGGCATCCCGGTGGTGTATGAGCAGAGCTGAGAGCTTGGCTAC 906
Qy 221 SerValThrAlaAlaLysGlnGlyIlePheGlnValValIleSerArgIleCysMetAla 240
Db 907 TCCTGTGACTGCAGCAAGCAGGAGATCTTCAGGTGATTTCAAAATCTGCATGGCG 966
Qy 241 ILleProAlaMetAlaIleProProLeuIleMetAspThrLeuGlyLysAspPheLeu 260
Db 967 ATTCTGCTCCATGGCGCATCCACCACTGATCATGTGACACTCTGGAGAAAGAACTTCTG 1026
Qy 261 Lys 261
Db 1027 AAG 1029

RESULT 4
US-10-014-338-1
; Sequence 1, Application US/10014338
; Publication No. US20030092614A1
; GENERAL INFORMATION:
; APPLICANT: Herath, et al.
; TITLE OF INVENTION: ADP1-41, A NOVEL PROTEIN ISOLATED FROM BRAIN TISSUE HOMOGINAT
; FILE OF INVENTION: USES THEREFOR
; FILE REFERENCE: 9195-077
; CURRENT APPLICATION NUMBER: US/10/014,338
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 10/014,338
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1121)..(1122)
; OTHER INFORMATION: where "n" is any nucleotide
; NAME/KEY: misc_feature
; LOCATION: (1125)..(1126)
; OTHER INFORMATION: where "n" is any nucleotide
```

[illegible][illegible]